Journal of
Clinical Psychology

OCTOBER 1955
Volume XI Number 4

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Editorial Office: 5 Pearl Street
Brandon, Vermont

JERRY W. CARTER, JR.
U. S. Public Health Service

HOWARD F. HUNT
University of Chicago

WILLIAM A. HUNT
Northwestern University

ELAINE F. KINDER
New York State Department of Mental Hygiene

C. M. LOUTTIT
Wayne University

ROBERT I. WATSON
Northwestern University

ELEANOR C. THORNE, *Business Manager*

Journal of Clinical Psychology
5 Pearl Street · · · Brandon, Vermont

READY NOW-COMPANION VOLUMES

Principles of Psychological Examining

by FREDERICK C. THORNE, M.D., PhD.

Editor, Journal of Clinical Psychology

- **The First Comprehensive Textbook of Applied Integrative Psychology.**
- **Systematic: Presenting a Complete System for the Differential Diagnosis of the Primary Etiologic Factors Which Actually Organize Socially Significant Behaviors.**
- **Complete: More Extensive Diagnostic Classifications of Etiologic Factors.**
- **Eclectic: A theoretical system is outlined for relating all pertinent theories and facts in basic science psychology and clinical psychological science.**
- **Advanced: More than a textbook of psychometrics. Tests and measures are presented only as they relate to more fundamental problems of psychological diagnosis.**

PRICE: $7.50

ORDER NOW

Principles of Personality Counseling

by FREDERICK C. THORNE, M.D., PhD.

Editor, Journal of Clinical Psychology

- **Systematic: The first comprehensive system based solidly on basic psychological science.**
- **Eclectic: Discusses all known methods with reference to their indications and contra-indications.**
- **Complete: For all teaching purposes.**
- **Advanced: A completely new orientation to diagnosis, counseling and psychotherapy.**

PRICE: $6.00

JOURNAL OF CLINICAL PSYCHOLOGY

5 Pearl Street, Brandon, Vermont

JOURNAL OF CLINICAL PSYCHOLOGY

VOL. XI OCTOBER 1955 No. 4

CONTENTS

Page

The Chief Invariant Psychological and Psycho-Physical Functional Unities Found by P-Technique. Raymond B. Cattell, University of Illinois 319

An Investigation of an Abnormality of Cognitive Function in a Co-operative Young Psychotic: An Example of the Application of Experimental Method to the Single Case. M. B. Shapiro and E. H. Nelson, Maudsley Hospital, London 344

Predictions by Means of Two Projective Tests of Personality Evaluations Made by Peers. Herbert Zimmer, Maxwell Air Force Base 352

Directive Therapy with Female Juvenile Delinquents. Sidney Kosofsky, N. J. State Home for Girls, Trenton, N. J. 357

The Discriminant Value of Minnesota Multiphasic Personality Inventory Items in Personality Disorders. Reuben S. Horlick, Walter Reed Army Medical Center, Washington, D. C. 362

Studies on the Physiology of Awareness: The Differential Influence of Color on Capillary Blood-Oxygen Saturation. J. W. Lovett Doust, University of Toronto, and Robert A. Schneider, University of Oklahoma School of Medicine 366

Relation between Histamine Tolerance, Visual Autokinesis, Rorschach Human Movement, and Figure Drawing. Stanley B. Graham, Creedmoor (N. Y.) Institute for Psychobiologic Studies 370

The Use of the MMPI in Predicting Irregular Discharge Among Tuberculosis Patients. Robert Calden and John R. Thurston, VA Hospital, Madison, Wisconsin, and Barbara M. Stewart and Shalom E. Vineberg, VA Hospital, San Fernando, Calif. 374

A Comparison of Right Hemiplegics and Left Hemiplegics in Visual Perception. Nancy E. Wood, Cleveland Hearing and Speech Center .. 378

The Patient Speaks: Psychologist-Patient Relations. Alvin R. Howard, Veterans Administration Hospital, Chillicothe, Ohio 381

Sociometric Choice and the Szondi Test. David Pearl, Veterans Administration Hospital, Battle Creek, Michigan and Durand Jacobs, Veterans Administration Hospital, Marion, Indiana 385

Teeth-Grinding: Some Relationships to Anxiety, Hostility and Hyperactivity. Francis F. Vernallis, Veterans Administration, Pittsburgh, Penna. 389

A Study of Ideal Psychological Adjustment. Anders Sweetland and George Frank, Florida State University 391

Goal Setting Behavior, Expressed Feelings of Adequacy, and the Correspondence between the Perceived and Ideal Self. Louis T. Lepine, Catholic University of America, and Bernard Chodorkoff, Veterans Administration Hospital, Dearborn, Michigan 395

A Comparison between the Peace Time Psychiatric Casualty Rates of Parachutists and Non-Parachutists. Charles Windle, Ft. Benning, Georgia and Harold E. Parker, 82nd Airborne Division 398

(Over)

CONTENTS—Continued

Rorschach Scorer Reliability. Richard H. Dana, St. Louis State Hospital 401

The Performance of Hospitalized Psychiatric Patients in the Ego-Strength Scale of the MMPI. Herbert Quay, Milledgeville (Georgia) State Hospital 404

Use of Eye Movements as a Response Indicator in Testing the Physically Handicapped. Donald E. Wamba, Public Schools, Corona, California and Stanley S. Marzolf, Illinois State Normal University 405

The Effect of Prefontal Lobotomy Upon Abstract Behavior. Albert R. Maslow, Catholic University of America 407

A Follow-Up Item Analysis of the Taylor Anxiety Scale. Arnold H. Buss, Carter Memorial Hospital, Indianapolis, Indiana 409

The Reliability of Wittenborn's Scale for Rating Currently Discernible Psychopathology. Shirley S. Pumroy, University of Washington, and William S. Kogan, Veterans Administration Hospital, Seattle, Washington 411

A Reliability Study of a Short Form of the WISC Vocabulary Subtest. Renate G. Armstrong, East Moline State Hospital, Illinois 413

Note on Hypothetical Situations in Personality Appraisal. Herman Feifel, VA Mental Hygiene Clinic, Los Angeles, California 415

The Development of Sexual Differentiation on the Draw-A-Person Test. Clifford H. Swensen and Kenneth R. Newton, University of Tennessee 417

Editorial Opinion 420

Book Notices 421

Subject Index, Volume XI, 1955 423

Author Index, Volume XI, 1955 427

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science. Special sections in each issue will be devoted to news, comments, and letters to the editors on matters pertinent to the profession. Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

ORIGINAL ARTICLES are published only with the understanding that they are contributed exclusively to the *Journal of Clinical Psychology*. The publishers are not responsible for statements made or opinions expressed by contributors in articles printed in these columns. Copyrights cover publication in the *Journal of Clinical Psychology* and articles may not be reproduced without permission of the publishers.

THE STYLE of all manuscripts including bibliographies must be in accordance with the instructions given in: Anderson, J. E., and Valentine, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.*, 1944, 41, 345-376. References should be arranged alphabetically according to the directions given above. Full address of the author should appear somewhere in the article, preferably at the end.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscribers' expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Rutland, Vermont, by Frederick C. Thorne, M.D. Subscription price $7.50; Canadian subscriptions $8.00; foreign subscriptions $8.50 including postage; single copies $2.00. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post-office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1955, Frederick C. Thorne, M. D.

THE CHIEF INVARIANT PSYCHOLOGICAL AND PSYCHO-PHYSICAL FUNCTIONAL UNITIES FOUND BY P-TECHNIQUE

RAYMOND B. CATTELL

Laboratory of Personality Assessment, University of Illinois

1. Levels of Concept Formation Regarding Functional Unities.

In the interests of effective design of current statistical and clinical experiments on personality in relation to physiological function it is desirable to take stock, at this research juncture, of the hypotheses brought out by P-technique investigations, even though the latter are yet few. P-technique has in many ways a claim to being the essence of the clinical method—raised to a new precision and potency. For it preserves the holistic approach while introducing measurement. In fact it takes measures simultaneously on many representative variables, and studies their coordinated changes with time, circumstance and, if necessary, also with controlled experimental influences. The peculiar potency of the method, however, is that, after so recording the course of each variable over time, it *sorts out the functional unities* behind the manifest variables, by the statistical processes of factor analysis[13] in a way which transcends the powers either of clinical method or of controlled univariate experiment.

Most theoreticians foresee considerable scope for the method in clinical psychology, both for checking on basic psychodynamic or psychophysiological hypotheses[12, 22] and for freeing the method itself, by varied use, of unresolved imperfections[13]. Nevertheless, such usage has scarcely begun, probably through dearth of background training and facilities[16]. There is every prospect at the moment that these restrictions on research are about to be remedied. Electronic computers, for example, are making the forbidding factorial labor a matter of the past. And it is for this reason that an orientation of current research effort, by the best possible interpretation of existing pioneer studies, is particularly important.

The formation of hypotheses about functional unities in personality has proceeded so far by methods at two distinct levels of penetration. At what may be called the "clinical hunch" level, certain fragmentary observations are integrated by an inspiration into a single concept such as "the super-ego", the "authoritarian personality" or "rigidity". Unfortunately, as well illustrated in the last two named, such hunches shape themselves more easily and extensively in the armchair than the clinic, and are apt to be a projection of the experimenter's personal mode of thinking rather than the sensitive reception of a pattern in nature. At the second level, as illustrated by, among other techniques, the device of P-technique, the experimenter first accumulates fairly complex observations, unmotivated by those traditional stereotypes which he shares with the man in the street or the fashionable professional theory of the moment. By the factor analytic method he is made aware of patterns that are more refined, more complex, more reliable, and more quantitatively defined. With his eyes opened to virtually new information he can discard the cruder stereotypes and thus begin his hypothesis formation at what might be called the "instrument revealed" level.

Psychology, and the kind of physiology adopted by many psychologists, has for a generation cluttered its theory with polysyllabic elaborations of popular concepts. Anything that might disturb those simple verbal integrations is often unwelcome. Clinicians have been heard to say of the dozen or more patterns revealed by P-technique that they are incomprehensible—and it is implied that they must, therefore, be wrong. But the aim of science is to find the new, and to the scientific psychologist these empirical regularities at a new level of precision are not viewed as an unfortunate distraction from time worn phrases—or as a veritable threat—but rather a challenging possibility of progressing to more effective clinical concepts. The aim of this paper is to concentrate on the six to a dozen patterns that have been confirmed by two or more researches and briefly to indicate hypotheses about these sufficiently founded patterns, as a basis for still a more precise research phase.

2. Objectives of Unitary Trait Matching

A word is necessary about the relation of unitary traits in P-technique to those in R-technique. Theoretically, the present writer has taken the position that any functional unity in nature should prove itself by appearing in both systems of covariation. But in practice there are limits of measurement reliability and inaccuracies of calculation which forbid us to apply this last test too rigorously. For in the first place, a pattern of covariation, such as that produced by the factor of general mental capacity, may reveal itself in individual differences, i.e., in R-technique, very readily, while the day to day variation of parts in one person may be so slight as to approach that zero value at which theoretically the pattern is incapable of detection by P-technique. On the other hand, the converse proposition, that some P-factors will fail to appear as R-factors should not hold. For any variation pattern *within* a person will appear in differences *between* persons, providing we freeze all persons at an instant, i.e., if we do not average measures made at many different times. Secondly, due to selection by other influences on some of the variables involved in a factor, the pattern, though influencing the same collection of variables as before, may produce a significantly different projection of loadings.

Present findings certainly fit these theoretical expectations: only *some* R-technique factors have appeared in P-technique, and these with some specific distortions. Moreover, the factors which appear in P-technique most *clearly* are those, such as surgency-desurgency [12] and the ergic drive patterns [14], but not such as intelligence or the acquired sentiment patterns, thus meeting the expectation that unitary traits psychologically known to fluctuate powerfully with internal and external stimulus conditions will appear most indubitably. In the present study we propose not only to match P-technique factors from one research to those of another, but also, and secondarily, to attempt matches with R-technique factors. In the latter we can fortunately build on the recent survey integrating evidence as to the most probable form of the principal invariant R-technique factors [16]. The present article may be considered the equivalent, in the P-technique realm, of that integrating R-technique survey, except that it is more tentative and indefinitive, due to the decidedly lesser development of the field.

3. Conditions of Matching

To date there are only seven P-technique studies [2, 18, 19, 20, 36, 42, 64] on human beings (and three on social groups) yet published, and we must restrict to five[1] of them because one deals with motivational variables only, and has no second study to confirm it, and another, although in the general personality realm, has no marker variables chosen to lock in with the other studies. These five studies have fortunately combined a fairly wide range of repeatable personality measures (initially those R-technique variables susceptible to repetition) along with physiological measures and records of everyday behavior. It must be admitted that the often heavy representation of physiological variables in these studies has been due as much to their seductive practicableness for repeated measurements as to the special fitness of the method for tying down psychophysiological functional connections.

In any factor analytic research, partly because of the complex consequences of error, but principally on account of the all too prevalent failure to rotate to adequate simple structure, there is no point in starting theoretical speculation on a factor until its pattern has been at least *once* independently reproduced. And even when, as is rare, experimenters have given definite attention to planting marker variables in their studies to ensure integration with other research, various experimental failures of variables themselves usually result in ultimate available markers being woefully few for the matching task. Such restrictions have reduced us in the end to seven

[1]Even so, Karvonen's study [42] is not exactly comparable with the others being a hybrid of R- and P-techniques. However, its correlations are of changes *within* persons and should yield essentially the same factors.

stable factor patterns, emerging from 43 phenomenal[11] factors, as will now be evident.

The procedures and standards in the following matching of factor patterns from study to study are essentially as in the parallel R-technique survey[16]. Each factor in one study has been systematically paired with every other factor in other studies. Except where mentioned it can be assumed that no alternative match of the factor loading pattern could claim any rivalry with that made here. In the tables the letter at the head of each column of loadings indicates the research (W for Williams[64], K for Karvonen[42], C1 and C2 for the two Cattell studies[18, 20] and H for Haverland[36]). The number labelled "size of sample" indicates the number of occasions the correlations are based upon—the time-and-circumstance sample—and the "factor number in the series" gives the order location in the series of the given study, as published. (The number has a minus beside it if the factor as published had to be reversed in direction). *The test variable titles have already been reversed in direction where necessary to agree with a negative loading.*

In general only loadings of .30 or more have been selected to mark a factor, but a variable so selected in one research has its loadings set down for the "matching" research even if it proves there to be low or of different sign, that the reader may perceive the degree of agreement. Disagreements of sign are emphasized by being put in square brackets. The reader should be reminded, however, that the degree of agreement is not appreciated from the tabled values alone, but only by remembering that there are also more variables than those shown *which agree by being essentially zero on both factors concerned, and are therefore not shown in the salient variables.* A "—" in a column does *not* indicate a zero loading, but rather that the test in question was absent altogether from the research, so that no matching figure could be expected.

In each table the variables occurring through *most* of the matched studies are first listed, and then, below the line, there follow the outstandingly high-loaded variables that are nevertheless not represented in more than one research. After the objective tests follow observer ratings and self ratings on everyday life behavior. The behavior to be rated was not always worded exactly the same in different studies and condensation of similar meanings has not been attempted where doubt exists. Thus each factor table may be read from top to bottom in roughly declining order of confirmation, magnitude and objectivity of elements in its pattern. The reader must be referred to the researches concerned for detailed description of the test procedures, though the meaning of variables will be brought out in discussions. Paralleling the R-technique procedure[16], we shall center the identification on a proposed *Universal Index* number, rather than a name. A "P" will be kept before these U. I. numbers to distinguish them from the R-technique series.

In the following section the listing will be confined to those twelve patterns, psychological, physiological or both, which seem best substantiated by present P-technique studies in general personality variables. There will also be listed, but without repeating the supporting factor tables from elsewhere, the eight patterns best substantiated in purely motivational dynamic variables. Universal Index numbers for reference will be assigned to seven of the former and eight of the latter. To place the cutting point between seven and eight in this series, in this order of diminishing degree of confirmation, is arbitrary, but, in the writer's judgment, appropriate, in terms of a conservative evaluation.

4. The Seven Confirmed, Indexed, General Personality Factors, with Somatic Associations

As Table 1 shows, the consistency of P. U. I. No 1 pattern, from study to study, is very high, and indeed the separate pairings brought together in the final matching here were first easily recognized individually, as the studies appeared. The ratings are also consistent, and, especially in C1, clearly give the picture of the H factor[12] of Adventurous Cyclothymia, (distinguishable from A and F because the latter occur in the same matrix, C1) first found in R-technique behavior *in situ*[11, 12, 15].

TABLE 1. FACTOR P. U. I. NO. 1: PARASYMPATHETIC RESILIENCE OR "ADVENTUROUS CYCLOTHYMIA" (*H* FACTOR)

Factor in given series	1	4	6	–2
Size of occasion sample	55	54	110	75
Particular research	C1	C2	W	H
Large % PGR deflections	73	68	50	84
High initial PGR resistance	—	40	45	67
Low glucose concentration in blood	—	–53	–40	—
High PGR upward drift in relaxation	39	—	38	60
Short dark adaptation (or neg. of im.)	—	–38	–49	—
High ataxic sway suggestibility	75	—	15	—
High ratio emot./non-emot. recall	65	—	12	—
Dreams described lengthily	—	70	—	—
Low atmospheric temperature	—	—	—	–61
High volume of urine per day	—	—	60	—
Low cholinesterase in blood	—	—	–56	—
Small lag of flicker fusion thresholds	—	—	–50	—
(Low pulse rate	—	—	–29	—)
(Estimated basal metabolic rate*	—	—	21	—)
Observer Ratings				
Friendly vs. Shy	44	49	—	—
Cooperative vs. Obstructive	44	—	—	—
Self confident vs. Unsure	44	—	—	—
Outgoing emotions vs. Quiet tolerance	44	—	—	—
Vigorous vs. Languid	—	44	—	—
Self Ratings				
Self-sufficient	—	–35	—	—
Own fault if things wrong	—	—	–20	—
Adventurous, high sex and food interests	—	—	15	—

*Added as second highest loading of BMR in the H study, though below cutting point here.

In relating the present longitudinal patterns to R-technique or "individual difference" patterns one is greatly hampered by the absence of any systematic attempt, since the present writer's in 1946 (and less extensively in 1950) to reach integrated conclusions regarding the central tendency of the published patterns in the realm of *physiological* variables. Without space to discuss the impact of some later findings upon the exact form of the R-technique patterns as of 1946 we must seem perhaps dogmatic in saying that the present pattern best fits that labelled by the present writer in 1950 as "parasympathetic pattern".

It contains, in the recognized textbook vagotonic pattern [1, 27, 34, 44, 48, 62]: low cholinesterase, large PGR response, high volume of urine, low pulse rate and high PGR resistance, while, in relation to the 1950 pattern, it also contains small lag of flicker thresholds, large resistance rise in relaxation, and low blood sugar. The outgoing H+ factor characteristics found here are also in close agreement with the associations summarized for large PGR response [11, p. 445] and for Meeland's "ready pulse recovery" measure [45], for Darrow and Heath's pattern [11, p. 444], and Wenger's [62] and Freeman and Katzoff's pattern [31]. The three last, we now believe, were wrongly identified with the "general autonomic activity pattern" (see factor 9 below) of Herrington [37]. In labelling this the "parasympathetic" pattern, as in any labelling of physiological factors here, we are aware that it may not fit one or another of the conceptions of the parasympathetic pattern variously described by Alexander [1], Fulton [34], Morgan [48] and others. Some, for example Wenger [63], so describe sympathetic and parasympathetic that we should expect them to appear as a single bipolar factor. Our factor 5 below is hypothesized to be the sympathetic pattern, i.e., we have two factors, opposed in some variables, unrelated in others and even together in a few. However, we might expect the *factors* to have some negative cor-

relation and this is true of our (mean) factor, though, except in C2, it is not as large as their supposed mutual exclusion might require. But semantic problems are to be avoided, and what we are essentially hypothesizing is that *a pattern can be reinstated* in future research as here defined and symbolized by P. U. I. No. 1, regardless of whether this proves to be wholly a parasympathetic activity product, or freedom from fearfulness, or whatever other hypothesis we prefer. The effect we are claiming that Wenger's factor[63] splits into three—autonomic, sympathetic and parasympathetic.

Our essential hypothesis here is thus: (1) the fluctuation pattern is essentially the same as the individual difference temperament pattern temporarily labelled "adventurous cyclothymia" and symbolized by H[11, 12, 15, 17, 45] and; (2) all people some of the time and some people all of the time are at a higher level of "parasympathetic activity"[2], less exhausted by sympathetic activity or fear, more anabolic, and consequently with those greater reserves of emotional energy which constitute H+. The appreciable correlation of H+ "energy" with lower temperature suggests that reaction to weather is one influence in generating H+ behavior. Meeland found the H factor in R-technique associated with good general physical adjustment to stresses and with little response of pulse rate to such a trauma as prolonged physical constriction of the arm.

Although the association with behavior rating and physiological factors as seen in R-technique seems to us beyond cavil, the matching with R-technique objective personality test factors is not so obvious—either because the picture is blurred or because no objective factor yet discovered actually represents this factor. Using the Universal Index factor numbers[15]—henceforth abbreviated to U. I.—the closest similarity is with U. I. 17(–), Lack of Restraint, and next with U. I. 21, Energetic Decisiveness, but neither has sufficient variables in common to be convincing.

Without prolonged discussion a basic defense of this interpretation is not possible. For both the H personality factor, with its uninhibited ebullience, and the somatic parasympathetic pattern, with its general picture of unpunished autonomic reserves, still require clear formulation, respectively by psychologists and physiologists. But theinitial evidence on inheritance[17] favors the view that the physiological resilience is primary (in this case fluctuating for physiological reasons from day to day) and the psychological pattern of energy causally dependent thereon.

In Table 2, representing the next factor, the agreement of independent studies, on as many as eight variables, is again very striking, though in this case there *are* a few sign inversions in loading. It should be noted that F1 in [64] has been rotated to better simple structure and that this makes it more orthogonal and also puts the ketosteroids wholly in F2 [64]. On the other hand it is noteworthy: (a) that cholinesterase variability is made by our study to be associated with two distinct factors, psychologically almost opposite, and (2) that these factors (plus factor 9 below) account for practically all the variance in it ($.78^2 + .56^2$ accounts for 92%, though these figures are swollen by an accidental overestimate of communality). Whereas in Factor 1 the low cholinesterase went with more "manic" behavior here it goes with anxiety-depression (see footnote Table 2) and general slowing of most spontaneous mental activities (but not of will acts, for new mental sets, concentration and overcoming rigidity are high, agreeing with the generally admitted connection of desurgency with super ego function[29]. One cannot help connecting this cholinesterase correlation with the summary by Vorhaus and Kark[61] showing lower cholinesterase

[2]Lest the reader should think there is some contradiction of these results and those of Tod and Jones[59] and Richter and Lee[52] who find severe anxiety (usually with some depression) associated with an *increase* of serum cholinesterase, the resolution possible through the factor analytic concepts should be indicated. For in Factor 1 above, increased cholinesterase is associated with inhibition, but in Factor 2, which follows, it is not. The observed correlation, by investigators using the above simple correlation methods, will depend upon which of these factors was predominant in their measure of "anxiety". Clearer results would presumably be obtained in this field if measures for the distinct forms of inhibition in what we call factors F- (desurgency) and H- (schizothymia) were employed, instead of estimates according to the popular term "anxiety".

TABLE 2. FACTOR P. U. I. 2: DESURGENCY VS. SURGENCY (F (–) FACTOR).

Factor in series	2	–7	1	3–
Size of occasion sample	55	54	110	75
Research	C1	C2	W	H
Long reaction time	75	35	67	21
Slow reversible perspective	–43	—	–61	(05)
Alkalinity of saliva (High pH)	56	33	10	—
Poor memorization	–45	00	–61	–13
Few figures added correctly in two mins.	—	—	–61	–49
Low fluency association	–21	(31)	–38	–40
Low rigidity	–35	–04	(29)	–03
Early in subject's day	02	–06*	–10	–55
Low sway suggestibility	–02	—	–36	—
Low cholinesterase in serum	—	—	–78	—
Low body temperature	—	—	—	–70
Small work upset by sound distraction	—	—	–66	—
Little concern with detail	—	—	–66	—
Slow tempo	—	—	–66	—
Brief, laconic, dream recall	—	–48	—	—
Low atmospheric temperature earlier	—	—	—	–47
High white corpuscle count	—	46	—	—
Good ability adopt new mental set	—	42	—	—
Unhappy feeling tone in relating dreams	—	37	—	—
Depressed, worried vs. Cheerful, talkative	37	22	50*	—
Obstructive vs. Cooperative	37	—	—	—
Restless emotional vs. Quiet	37	57	—	—
Submissive vs. Self confident	32	—	60*	—
Daydreaming	32	—	—	—
Shy, self conscious	—	—	80	—
Self Ratings:				
Feelings easily hurt, depressed, lonely	—	—	80	—
Felt above average in perseverance	—	—	80	—
Kept to one track in conversation	—	—	66	—
High frequency awareness stomach pains (ulcer case)	—	54	—	—
Felt poor control of memory	—	—	90	—

*Corrections have been made in this row. With the re-rotation of W's F1 the behavior ratings could only be approximately re-determined, by projecting the correlations on the factor plot and shifting them with the other variables.

**Original data not available to permit check on uncertain sign in this correlation.

concentrations with liver disease, malnutrition, acute infectious conditions and anemia. The classical description of the mental condition associated with mild liver disorder is indeed what we would now call desurgency. Defective production of cholinesterase by the liver may thus be one causal condition of desurgency, and the correlation with early hour of day may be similarly mediated.

The objective test pattern of P. U. I. 2 is well matched in R-technique studies, using identical tests, namely by U. I. 22: Nervous, Alert, Reactivity (in reverse). It matches on long reaction time, slow reversible perspective, alkalinity of saliva, low fluency, low body sway and low rigidity. Meeland found a similar factor (F9 in [43]) loading long reaction time, slow reversible perspective and inability to solve riddles (intelligence held constant) the sense of which agrees well, though he found no relation to desurgency. U. I. 21(–), however has some similarities. From the earlier studies, and consistently in all, the test pattern has tied up with the rating pattern of Desurgency vs. Surgency, or F[18]. This is generally peculiarly difficult to distinguish, in rating rotations, from H, but it is more confidently identified here because these factors appeared simultaneously in the same study (C1).

It should be noted that the Haverland research is a noticeably poorer match than the others. This is *generally* true of Haverland's and Karvonen's studies—be-

cause they are differently oriented. The former repeats the testing twice and generally oftener on *one* day, thus bringing in far more diurnal time range than the other studies, and Karvonen's also concentrates on diurnal variation. Haverland's factor is considerably loaded in time of day (the only one that is), a person having a rising temperature and becoming more surgent as the day wears on(apart naturally from the simultaneous effect of distinct fatigue factors). But this diurnal surgency trend has slightly different pattern emphases than have the larger surgency changes from day to day, and its shape may be partly a peculiarity of the S concerned, who was a "late starter", anxious and sober (and perhaps "liverish") as he faced the day's work. Best and Taylor [4, p. 304] state that the cause of the diurnal metabolism swing is unknown. These results suggest the metabolism rise is associated with increased speed and fluency and the discharge of anxiety in action.

Our hypothesis for the Surgency-Desurgency factor in individual differences (R-technique) is that it represents the general level of cortical inhibition, associated with past punishment, cultural complexity, and demands of the self sentiment (and therefore including much that lies in the superego concept). It is to be expected that this same pattern will vary from day to day with different provocations of guilt, aspiration, etc., and perhaps with internal metabolic and intoxication changes. Some physiological sense can be made, as indicated above, of the association of depression with low cholinesterase, but the total pattern of physiological elements—low cholinesterase, low body temperature, alkalinity of saliva and high white corpuscle count—is one which we have not been able to find in physiological literature. The pattern is a challenge both to physiologists and psychologists. However, as a clear empirically established functional unity there would seem to be little doubt about it, either physiologically or as the surgency-desurgency found in R-technique. Our original hypothesis [11] that desurgency is a general inhibition or deficiency of mental energy—psychologically or physiologically maintained, and subject to fluctuation from either influence—seems well supported here and poses a question for crucial research on mechanisms, in both inter-and intra-personal differences.

The W study factor 4 matches in Table 3 here, and nowhere else; but it does not match well. In all four studies this is the principal factor associated with hour of day[3] (the only other factor, with slighter diurnal associations, having just been discussed above). It has been labelled "general fatigue" because it loads variables such as reversible perspective, upward drift on skin resistance when relaxing, flicker fusion, and lengthening of prewarned reaction time, which have been found or accepted in various special studies [8, 9, 26, 36, 66] as indicators of fatigue. In K's study in Finland, there is no doubt of the fatigue connotation, since his subjects were in a prolonged tree-felling competition. It is ironical that Haverland's study [36] specifically directed to fatigue variables [5, 8, 26, 42] never found this factor, but perhaps his high density of such variables forced general fatigue into the position of a second-order factor, for which his inter-factor correlation matrix remained to be examined.

Physiologically the pattern is clearly that of Selye's [64] late reaction and counter shock (G. A. S. stage two) comprising "intense neutrophilic leukocytosis, lymphopenia and eosinopenia". [42] Karvonen's differences from our pattern could be partly due to his small sample and partly due to simple "factorization of increments" [13] but, at least as far as the erythrocyte difference is concerned (the only large one) it is clearly due to his having used heavy *muscular* fatigue. For he points out that Brown [6], Dalton and Selye [23], Gurejew and Karvonen [42] and others [28] have found lowering of erythrocytes markedly in prolonged muscular fatigue but scarcely in other work. Our figures are presumably a compromise between the normal tendency for erythrocyte count to increase during the day, and the reduction of erythrocytes by muscular fatigue. Possibly the upward erythrocyte trend is associated with the

[3] In C2, hour of day was not included, but "frequency of recorded urinations" happens in other studies to correlate most among available variables with time of day, and this occurs in C2, with the only appreciable loading, in F6.

TABLE 3. FACTOR P. U. I. 3: GENERAL DIURNAL FATIGUE

Factor in series	3	6	4*	1
Size of occasion sample	55	54	110	60
Research	C1	C2	W	K
Hours after waking (or of day)	71	—	76	High positive
High ratio warned/unwarned reaction time	34	47	14	—
High calcium in serum	—	42	18	—
Recorded urinations	—	41	60	—
Low % lymphocytes	—	-36	(09)	-62
High % neutrophils (seg)	—	28	(-15)	71
Rapid reversible perspective	56	—	(-09)	—
Big upward drift or high original level of resistance in PGR	40	40	(-01)	—
High white corpuscle total	—	35	37	—
Low eosinophils	—	—	(37)	-64
Low erythrocytes	—	-01	(02)	-65
Happy feeling tone at end of dreams	—	-70	—	—
High frequency of laughter in interview	—	51	—	—

*On turning to the original published studies the reader will find slight changes of rotation on a few factors. The decision to integrate the P-technique studies caused the writer to return to the individual studies and bring rotations to a still clearer simple structure wherever necessary. The improvements affect only one factor in ten, as follows: In C1, none; in C2, none; in W, $F'_1 = -F_1 - 1.3F_2$, $F'_2 = F_2 - .50F_1$, $F'_4 = F_4 - 30F_1 + 10F_3 + 10F_6$; $F'_6 = F_6 + 20F_8$; $F'_7 = F_7 + 25F_4 + 10F_5 - 10F_{13}$; $F'_8 = F_8 - 25F_6$: in K the following, $F'_1 = -F_1 - 1.00F_2 + 20F_3$, $F'_2 = F_2 - .80F_1 - 20F_3$, $F'_3 = F_3 + 1.00F_1 + 20F_2$. (This includes, after correspondence with Dr. Karvonen, the extraction of a small 3rd factor (mean loading .05) from his data). In H, $F'_1 = F_1 + 70F_9 + 15F_4 + 10F_7 + 10F_2 + 15F_5 - 15F_6$ and $F'_9 = F_9 - 35F_1 - 77F_2 + 15F_4 + 15F_8 + 15F_5 + 10F_7 + 10F_{11}$. These rotations improved simple structure and it is noteworthy that in W1 and 2 they remove a confusion, making cholinesterase the essence of 1 and ketosteroids the essence of 2.

metabolic rate increase in the daily trend in P. U. I. 2 above. That this factor is not a vagotonia factor is shown by simultaneous appearance of the latter in C2.

Just how this "general fatigue" or "exposure to waking effort" factor will relate to other, more specific fatigues must be left to decision by factorial studies like Haverland's on a large scale. Study of lower correlations in our data will suggest further possible tests of the present fatigue factor, e.g., frequency of small, "spontaneous" PGR deflections. It may also indicate variables best suited to separate the other time-of-day-related factor ("warming up" or surgency) from the general fatigue factor. The latter is certainly well pegged down by the neutrophil, lymphocyte, eosinophil and perhaps serum calcium changes not present in the former. It is notable that no *personality ratings* or R-technique patterns associate with this factor. Since fatigue, like death and taxes, comes to all mankind we should indeed doubt our technique if prominent personality rating associations appeared here or if a decided R-technique factor could be matched.

The agreement of the next pattern, in Table 4, over four studies, is striking, the only sign disagreement on a significant value being on study C1 where this particular value was approximate [18]. Both tests and ratings give the psychological meaning of high emotional stability and will control[4], and indeed the pattern has earlier been identified with the rating Factor C [12], stability or ego strength, while being distinguished from G (with which it might otherwise be confused) through the G pattern occurring also in the same study, independently.

[4]The H study, probably for reasons already mentioned, is a bit out of line with the others, notably in having this factor significantly correlated with time trend. This could mean that practice improves the same performances as are susceptible to improvements by willed effort, at any given moment, or, alternatively, that this particular subject was in fact in process of a long term "therapeutic" trend to higher C factor (stability). The latter, from our knowledge of the case, is quite likely.

TABLE 4. FACTOR P. U. I. 4: GENERAL ADAPTATION SYNDROME ONE OR CONTROL (EGO STRENGTH) (C FACTOR)

Factor in series	4	-3	2	1	
Size of occasion sample	55	54	110	75	
Research	C1	C2	W	H	
Low fluency (verbal and drawing)	-55	-14	-29	-44	(Mean of 3 tests)
Low rigidity	-51	-09	-31	-42	
Small myokinesis	-40	-36	—	—	
Low lymphocyte count	—	-42	-22	—	
Low sway suggestibility	-35	—	-11	—	
Good memory, commit, and recall	(-40)	36	44	31	(Mean of 2 tests)
Small PGR deflections	-32	(05)	-05	-66	
Slow reversible perspective	-09	—	-22	-05	
High 17 - Ketosteroids	—	—	97	—	
Low ability to get new mental set	—	67	—	—	
Session number	—	—	—	49	
Effective multiplying	—	—	—	40	
Little time spent on preference rating	—	—	—	-35	
Little time spent on cards (interest)	—	—	—	-59	
High C. M. S. total score	—	—	—	46	
Few C. M. S. slanting lines crossed	—	—	—	-39	
Few circles used in excess	—	—	—	-33	
High cholinesterase	—	—	38	—	
Observer Ratings:					
Stable and calm emotionality	50	51	50	—	
Self-sufficient vs. Attention getting	49	38	—	—	
Not daydreaming	-47	—	—	—	
Solemn	36	—	—	—	
Adventurous (Enterprising)	—	—	30*	—	
Inhibited vs. Sociable	—	38	—	—	
Not emotionally insistent	—	51	—	—	
Self Ratings:					
Not excited and rattled	—	—	50*	—	
Felt good control memory	—	—	30*	—	

*Approximately estimated after rotation.

Some of the elements of the pattern have been observed together in pairs before, in univariate studies, notably Freeman and Elmadjian's observation of low lymphocyte with higher glucose[31, 32] and Forbes[30], Richter's[52] and Tod's[59] of high ketosteroid with high cholinesterase. (They note also some lowered alkalinity of saliva—loaded only a mean of .10 in our studies, below the cutting point in the table). In R-technique patterns, factors U. I. 19, 21, 26 and 29 all have some resemblance, and there is some real rivalry of 26 and 29, but 26 (entitled "control") agrees most conclusively in showing low fluency, low rigidity, good memorizing and high C. M. S. total score.

As to interpretation of these substantially combined psychological and physiological manifestations we must set aside such simple notions as a physiological reaction to infection, or the ready resort of many psychiatrists and physiologists to "a state of anxiety". (How ubiquitous can "anxiety" or "tension" become and remain meaningful?) The striking physiological association here is a loading of almost unity in adrenal corticosteroids, and this steroid output loads no other factor. It would seem that we have indeed separated out factor analytically the stress reaction of Selye, with its "adreno-cortical stimulation, thymo-lymphatic atrophy and gastro-intestinal ulcers"[54, p. 27] (except that the last would not have had time to appear) and what is perhaps of still more technical interest, the factor analytic method has separated this from the somewhat similar but distinct stress pattern which we have labelled the fatigue pattern (possibly G. A. S. stage two), in factor 3 above.

But what is the "stressor" in this case? And is the high C-factor, self control and steady effort a compensatory psychological response to waning physiological resources produced by physiological stress, or is it psychologically causal to the physiological pattern? If the 0.97 figure were the mean of several studies we should have no alternative but to accept the first view, i.e., to consider this higher loaded physiological variable the cause or factor. Then we could try either the above hypothesis of the "sobering-up effect" of expended resources of steroids or take an alternative physiological causation and suppose high drive and stability springs from high physiological stimulation. The latter would presumably be identified primarily with powerful ACTH output, which would, in our data, be showing its effects by beginning to encourage high flow of corticosteroids. The purely clinical evidence on steroid levels in mental patients and on effects of therapy with these hormones is highly contradictory [7, 51, 65, 56] perhaps for reasons discussed below. But there is a beginning of consistency and confirmation from the Hoagland [39] and Pincus [60] findings of diminished steroid output and diminished adreno-cortical sensitivity to ACTH in schizophrenics, and in Reis [61] and others' [65] finding of improvements in certain kinds of depression through administering ketosteroids or ACTH.

On either our own or the clinical evidence we favor the second major hypothesis here, namely, that the psychological demands in "emotional control" and persistent effort (C factor) constitute a "stressor" which produces, presumably via the pituitary, higher steroid output. If, however, as in the clinical cases, the adrenals are chronically incapable of response, the individual completely abandons such efforts at will control (whether this means relapse into depressive, schizophrenic-manic or paranoid expressions is determined by secondary factors). In short, this would add to Selye's battery of "stressors" the will or ego itself. The "self-sentiment in action" [12] which we believe C factor to be, is presumably triggered by psychological stimuli —environmental demands on the self-concept—and this produces the physiological stress. There is ample independent evidence [19] that the self-sentiment varies in level of functional activity from day to day.

Browne's observation [7] that in women "going to bargain sales, adrenal cortical output went up definitely" supports clinically our statistical pattern of higher steroids with the stronger exercise of will qualities of perseverance and persistence. Whatever the additional evidence needed for this interpretation, the initial evidence that the psychological pattern is C factor in ratings, and U. I. 26, with the similar meaning "self control" in objective tests, is unequivocal. The only suggestion of any deviation from these "classical" R-technique concepts is, incidentally, a certain persistent "emotional insistence" and abundant spontaneity (evident particularly in lesser loaded C1 and C2 variables) at the general emotionality or "uncontrol" pole; but some R-technique studies have also indicated the C- is not all inactivity and demoralization. Our hypothesis is then that this control itself, over impulse and fatigue, is purchased physiologically, as well as psychologically, at cost of stress. If correct, it may help explain why business executives, without physical labor, have a higher death rate from heart disease than farm laborers.

Actual agreement of patterns is still statistically satisfactory in Table 5. Only two disagreements of sign exist, each at non-significant loading levels, out of 21 tested associations. Nevertheless, the Karvonen factor cannot be taken as much support for the pattern, since it has only three variables in it. (In univariate studies this would nevertheless be considered three confirmations of an hypothesis!) Physiologically the connection of high lymphocytes, and sometimes high calcium, with high glucose has been sporadically noted before by Dunbar [27], and Kretschmer and Kruger [64]. Tomasson [58] and Glaser [35] note manic and "nervous" individuals to be high in calcium. The pattern is in fact one of being "pressed" in Cannon's sense [41] and fits the classical sympathetic response in high pulse rate, blood sugar, large resistance rise in relaxation (low basic resistance), and high erythrocyte count. The common appearance of high glucose and high erythrocyte count would, of course, indicate adrenal action. The decided low neutrophil count has not been mentioned

TABLE 5. FACTOR P. U. I. 5: ADRENERGIC, CYCLOTHYMIC STATE (A FACTOR).

Factor number in series	5	7	3	4
Size of occasion sample	54	110	60	75
Research	C2	W	K	H
High % lymphocytes	48	48	49	—
Small % neutrophils	-30	-67	-52	—
High glucose in blood	34	54	—	—
Poor memorizing	-58	-23	—	-38
High ratio warned/unwarned reaction time	55	(-03)	—	14
High calcium in serum	08	46	—	—
Large PGR upward drift in relaxation	—	16	—	22
High erythrocyte count	(-05)	09	44	—
Dreams lengthily described	69	—	—	—
Small difference flicker fusion thresholds	—	-68	—	—
Many slanting lines on C. M. S.	—	—	—	50
Many circles used in C. M. S.	—	—	—	45
Rapid pulse rate	—	—	—	38
Low oscillation	—	—	—	-38
Preference for mental over physical work	—	—	—	37
Observer Ratings:				
Cooperative, not obstructive	35	—	—	—
Cheerful, not lonely or over-sensitive	—	25	—	—
Absorbed, not distractible	37	27	—	—
Self Ratings:				
Inclined to stuttering	—	47	—	—
Sociable, not shy	—	30	—	—
Self confident	—	27	—	—
Memory under good control	—	25	—	—
Not easily startled or distracted	—	21	—	—

in this connection but is strongly indicated here to be part of the pattern. (This decline should account for the high lymphocyte percentage.) High lymphocyte count also occurs in the composite factor[11], T(SF) I from Darling[24], Herrington[37] and Wenger[62], but without some other, lesser associates indicated in the non-factor or unrotated factor studies, e.g., pupil dilation, systolic blood pressure[31], metabolic rate; but these, as far as correlation analysis evidence goes, *could* be a distinct but correlated factor[16].

Psychologically, the pattern is very close to that of cyclothymia (cooperative, friendly, day-dreaming (phantasying), vigorous, "jumpy") as found in P-technique ratings before [18], except for the "jumpy" rating. But the difficulty in relating sympathetic excitation level to cyclothymic state is that the former has usually been considered a reaction to "threatened" psychological states and indeed there are some features here, e.g., stuttering, panic errors in C. M. S., which fit a hypomanic rather than merely cyclothymic picture. However, until further precision is gained, the best contingent hypothesis is that this is the adrenergic "high drive" response pattern found in a person sensitively interacting with his environment and that it goes with the rating pattern scarcely distinguishable from that called A, or Cyclothymia. The clinical picture of adrenal *deficiency*, in Addison's disease, with its central pattern of apathy, negativism, and seclusiveness certainly agrees well with the A- pole here, suggesting that within normal ranges the behavior trends with adrenal sufficiency-insufficiency remains the same as indicated by abnormal extremes.

In R-technique, *objective* personality tests, a good match is found with U.I. 19(-)[16] or Casualness vs. Critical Exactness, resting upon: Many slanting lines on C. M. S., unrestrained C. M. S. circles; large PGR upward drift and the general sense of disregard for cautious exactness. The basic hypothesis resulting is that, in interpersonal terms, the schizothyme individual somehow protects himself from appreciable sympathetic reactiveness, whereas the cyclothyme does not, and, intro-person-

ally, that any individual in his more cyclothyme moods shows a higher sympathetic excitation level (when we define the physiological sympathetic pattern essentially as here.)

In all three studies one factor was found which loaded conspicuously the order of experimental sessions as shown in Table 6. This can be initially regarded statistically as a "partialling out" factor for all that is "artificially" associated within the experiment with *sequence*, namely, familiarity, learning and/or boredom and refractoriness to repeated experiment. Some variables here fit this explanation well, notably C. M. S. skills, rigidity decline and threshold changes (all learning); but others point to a time trend having nothing to do with the experimental experience as such. For example, the heavy representation of physiological variables is difficult to understand except by extreme reliance on supposed physiological effects of presumed boredom.

TABLE 6. FACTOR P. U. I. 6: LONG TERM PHYSIOLOGICAL "WARM SEASON ADJUSTMENT" TREND—PLUS PRACTICE.

Factor number in series	8	3	–9
Size of occasion sample	54	110	75
Particular research	C2	W	H
High calcium in serum	37	71	—
High lymphocyte count	46	39	—
Order in session	60	–52*	47
High verbal fluency	42	–39*	(–9)
Low blood pressure (diastolic)	—	–55	–31
Low pulse rate	—	–26	–13
Low rigidity	–02	00	–46
High pulse pressure	—	37	05
Low slanting line score C. M. S.	—	—	–51
Low circle mass score C. M. S.	—	—	–37
High C. M. S. total score	—	—	34
High sensitivity tactual threshold	—	—	35
Large diff. flicker thresholds	—	–45*	—
Low % eosinophils	—	–31	—
Unhappy feeling tone in dreams	57	—	—
Dreams reported incomplete	45	—	—
Little time spent on postcards	—	—	–33
Observer Ratings:			
Self-assertive	—	20	—
Self Ratings:			
Distractible (Bored?)	44	—	—

*See note in text

Serious consideration should therefore be given to the alternate possibility that there exists a factor which fluctuates so slowly—perhaps in one direction for a season, a year, or more—that the usual two or three month P-technique study is likely to catch it moving in only one direction. Support for this comes from the peculiarity that the starred variables in study W are actually loaded in the reverse direction to the other studies, suggesting the trend is there caught in the downward direction. (These variables are nevertheless not bracketted, as sign exceptions, because as far as the *learning* trend is concerned they are consistent in one direction, and as far as the *physiological variables* are concerned, they are consistent in the other. They fit, in other words, the hypothesis that the external trend can be going either way with respect to session order. In any case, all the non-practice variables (with possible exception of unhappy feeling tone in dreams and the rating variables, which could be boredom-revolt) are unrelated to *direction* of time and are physiological in nature. This possibility of there being two factors, the practice factor and the physiological

factor, could be tested further by including such hyperplane variables as could split the present factor into two, if it really is fissionable.

Meanwhile, we may attempt to deduce from the non-practice variables what this slow trend could be. Cureton has suggested that the physiological pattern corresponds to a factor representing fitness through physical exercise and which he and his students (see, e.g., Heusner [38]) have shown (at least as a correlation cluster) marked by low diastolic blood pressure, low pulse rate, high pulse pressure, high lymphocyte count and, possibly, high calcium. This meaning would accord with the present factor in that the better C. M. S. performance could also well be a function of fitness through exercise, and it is certain that our student subjects fluctuated a good deal in exposure to physical exercise over the semester in which testing occurred. On the other hand, an equally promising hypothesis is that it represents a seasonal weather adjustment! There is ample evidence in physiology (see Wolfe in [40, 53]) of discernible biochemical and weather adjustment trends, and Haverland's psychological data [36] showed unquestionable "weather factors". Sargent [53] points out that in the spring movement toward warmer weather there is a rise in calcium serum level and a drop in diastolic pressure and pulse rate. This hypothesis would require that the subjects in the C and H studies sat at the opposite phase of the year from that of the W study, and on consulting the records we find this is indeed the case. The C and H subjects sat from mid April until June and the W subject from November to March. Accordingly, and since C and H show the increase of calcium, etc., required by Sargent's generalizations, we favor the conclusion that this is a warm season adjustment factor.

TABLE 7. FACTOR P. U. I. 7: BASOPHIL-NEUTROPHIL PATTERN

Factor in series	12	2
Size of occasion sample	110	60
Research	W	K
Many staff nuclear neutrophils	37	61
Few basophil cells	–34	–32
Few eosinophils	–33	–09
High frequency juvenile neutrophils	—	66
High specific gravity urine	81	—
Small % poly-neutrophils	–40	—
High drawing fluency	36	—
High verbal fluency	31	—
Large diff. of flicker thresholds	33	—
Observer Ratings:		
Accommodating, communicative	34	—
Persevering, conscientious	29	—
Self Ratings:		
Steady, persistent	32	—
Emotional moods under control	32	—
Not anxious or worried, in good spirits	28	—

This pattern in Factor 7 (Table 7) is of interest as the only one with appreciable and confirmed loading on the basophil cells. The high specific gravity of urine has been associated, as low urine volume, with freedom from anxiety states by Mohr [46] and this agrees with our ratings. But again we recognize the existence of *varieties* of anxiety vs. non-anxiety, and in this case one is reminded of the 0 factor of free-floating anxiety in the questionnaire series [12], with its opposite of calm assurance and resourcefulness. Although there are such consistent psychological associations here, they have small loading and the pattern is therefore likely to be one primarily determined by physiological conditions.

5. Survey of Less Confirmed Patterns in Ratings, Objective Tests and Motivational Data

Of the 41 published objective general personality factors in seven studies, 25 appear as recurrences of the above seven patterns. Of the remaining 18, probably 6 more deserve tabulation, as 5 distinct factors, worthy of experimental and theoretical consideration, on grounds of substantial variance, clear structure, adequate sample and occasional R-technique resemblances. Except for a vagotonia-like factor remaining unmatched in C2, the 11 "odd factors" which will then remain are found in study W, which, due to communality excesses, probably extracted too many factors, so the reader is referred to these only where they are set out, as suggestive for further studies. Universal Index numbers will not be assigned to the five factors now to be listed, since they cannot be considered sufficiently confirmed, and they will have the temporary index P. F. (P-technique Factor, at phenomenal level).

Table 8. P- Factor 8: Untapped reserves, insensitive vs. overwrought.

Factor in series	1	−10
Size of occasion sample	54	75
Research	C2	H
Marked completeness of dreams	−92	—
Low speed of writing and multiplying	−70	(−10)
Low rating of funniness of jokes (Also recorded laughter)	−55	—
Low disposition rigidity (motor)	−54	(−17)
Long hours of sleep previous night	51	26
Large movements in myokinesis	44	—
Quick reaction time	−41	−14
High fluency of association	07	38
Low steadiness of hand	—	−42
Preference physical over mental work	—	29
Long endurance ergometer	—	29
Early in day	—	−21
Observer Ratings:		
Languid	−29	
Retiring	−27	
Attention getting	27	
Self Ratings:		
Low strength interest in social activities	−63	

The Haverland study factor most loaded in hours of sleep is a *possible* match here, as shown, though poor. Probably sleep records have been the least reliable data in these studies (for we did not realize how much training most subjects require), but this factor suggests long sleep is relevant enough to be accurately experimentally pursued. There is energy here, as shown by quick reaction time, fluency and low rigidity, but disinclination to use it. Perhaps the popular phrase "drugged with sleep" well describes the state caught here. It has the second highest diurnal time association, though still low enough ($r = .21$) to suggest that the condition tends to last through the day, i.e., as a state of being *well* rested, not merely *recently* rested. In R-technique factors it has some relation to U. I. 21 and 25 but most clearly to U. I. 23 or "neural reserves", which it matches in low rigidity, high fluency, short reaction time and high endurance. Eysenck [29] found U. I. 23 to be the principal factor distinguishing non-neurotics from neurotics. If the present match holds, it supports our contention [16] that U. I. 23(-) is not neurotic conflict as such, but a dimension of "shortage of neural reserves", or, in the present context, "overwroughtness".

Earlier it was suggested [20] that the pattern of Factor 9 represents some metabolism deficiency associated with paranoid-like behavior. Closer scrutiny in conjunction with factors P. U. I. 1 and 5, indicates that it may also be considered the ob-

verse view of what has previously been recognized as T(SF)I or "general autonomic activity"[11]. The factor so described by Herrington[37], Miller[45] and Wenger[63] has also typically contained pulse rate, PGR magnitude and body temperature. (The first two are correct on sign here –06, –21, and the last was not in the experiment.) Meeland has a similar R-technique factor (No. 2 in his series) loading low basal metabolism, high skin resistance and low body temperature[45]. Darrow and Heath[25] describe a roughly similar cluster associated with paranoia. Both of the last studies could almost as well be associated with P. U. I. 1 here, and the existence of some "cooperativeness" or admitted difficulty in separation of P. U. I. 1 and 9 must be considered, paralleling the closeness of factor H and L in ratings. The psychological associations are in agreement with Herrington's but the correlations are not as great as he indicated. In view of the few cases and contamination with other factors in earlier studies we would revise our 1950 conclusions that Herrington's "syndrome" is H factor and fit his "factor" to a smaller role than then (on his very few cases) suggested itself from his correlations. It seems that we must recognize this as a "general autonomic", blood pressure and basal metabolism factor; not affecting as many variables as previously supposed and essentially connected with an altruistic cyclothyme vs. paranoid schizothyme rating pattern, i.e., factor L[12], rather than H. The main contribution to the problem from present findings is therefore the separation of H and L, and the indication of high association of this metabolism pattern with serum cholinesterase (again presumably mediated by defective liver metabol ism). It is noteworthy that this factor, taken with P. U. I. 1 and 2 above, would account, as stated, for virtually all variance in serum cholinesterase.

As many clinicians have noted, many diverse debilitating conditions can bring on paranoid-like states. Our hypothesis is that this is a state of lower cerebral and general metabolism, triggered by physical autonomic conditions, and resulting in a shift on the L factor axis, toward a paranoid mood. In the opposite direction it is a factor of high general autonomic activity with friendly enthusiasm and desire to excell[11].

The concurrence of PGR response with dilation of pupil has been called by experimentalists, e.g., Woodworth[66], the "autonomic reflex", but our factorization

TABLE 9. P- FACTOR 9: AUTONOMIC INACTIVITY OR LOW METABOLISM WITH PARANOID MOOD (L FACTOR)

Factor in series	5
Size of occasion sample	110
Research	W
High cholinesterase in serum	78
Low pulse pressure	–71
Low basal metabolic (estimated)	–59
Low pH (acid) saliva	–42
High ratio emot./non-emot. recall	–37
High pH (alkaline) urine	32
Many staff neutrophils	30
High initial PGR resistance	25
(Slow reversible perspective*	–19)
Self and Observer Ratings:	
Shy, withdrawn	–30
Tense, worried, anxious	–23
Observer Rating:	
Submissive	–25
Self Rating:	
Fault of others when things wrong	21

*This is included, despite below cutting point, to check on Kris's finding[43] that speed of reversible perspective is correlated with metabolic rate.

(Table 10) shows that it is distinct from the three primary autonomic factors, (P. U. I, 1, 5 and 9), though it is "cooperative"[13] with them. In R-technique a fourth autonomic pattern is known—through Darling[24] (T(SF) IVa or Nervous Disposition in [11]) which matches this in pupil dilation, large PGR reaction, high sensitivity, low blood pressure and high alertness but disagrees on skin conductance and probably better matches P. F. 11. There is also objective test, R-technique, personality factor, U. I. 29, or "Ready Mobilization of Energy" which matches well on: High ratio of emotional recall, high skin resistance, low rigidity, low ratio regular/irregular reaction time, indifference to distraction, and large PGR deflections (but missing on flicker fusion frequency, memorizing and computation speed)[16]. The latter may be *long term* products of a high level on this temperament factor, and not responsive to rapid diurnal swings. Meeland's F-7 combines similarly high ratio of emotional recall, with large PGR deflection, and adds questionable preferences and slow speed of work. At present the pattern can be stated but not interpreted, except in general terms as some kind of heightened emotional liveliness and uncontrol.

TABLE 10. P- FACTOR 10: "AUTONOMIC REFLEX": NERVOUS DISPOSITION

Factor in series	8
Size of occasion sample	110
Research	W
High critical freq. flicker fusion	65
Large pupil diameter	56
High ratio emot./non-emot. recall	53
High initial PGR resistance	50
Low tolerance level electric shock	–45
Low specific gravity urine	–38
Large PGR deflections	34
Observer Rating:	
Adventurous, debonair	22
Self Rating:	
When wrong, mainly my fault	–25

No match in other P-technique series has been found for Factor 11, but its substantial variance, clear hyperplane and psychologically consistent nature justify recording it. An R-technique physiological cluster or factor is known through Darling[24], Herrington[37], Miller[46], Wenger[62, 63] and Cattell[9], (T(SF) III[11]) somewhat like the obverse of this factor, matching principally in low skin resistance, and which has been called a "hyperthyroid factor." However, basic metabolic rate is more central there than it can be demonstrated to be here. There is also some resemblance of this factor to objective personality test factors U. I. 17(–) and 21(–), but insufficient evidence to choose.

The greater loading here than elsewhere (except C factor) in conscientiousness and obligingness raises the question whether this factor can be considered the G or functioning super ego pattern known in terms of R-technique ratings and already discovered in P-technique[18] as a loading pattern on: persevering, steady, self-sufficient and cautious shyness. As we do not know how much the rating loadings here need to be corrected for attenuation, a decision on identity cannot be made; but we can conclude (1) that there exists a fluctuating pattern in ratings very similar to R-technique G and (2) that it has an appreciable relation to the physiological pattern here. If the original notion of a hyperthyroid factor (T(SF) III in[11]) is substantiated by additional variables, and identified as this factor, then we should have the rather strange finding that super ego manifestations decrease slightly in the normal individual in hyperthyroid states. Perhaps lack of super ego expression generates guilt and anxiety, stimulating thyroid action. In the typical myxoedemic syndrome of

TABLE 11. P– FACTOR 11. SUB-THYROID OR FATIGUE AND SUPEREGO (G FACTOR).

Factor in series	9
Size of occasion sample	110
Research	W
Low pH (acid) saliva	–80
Poor memorizing	–63
Short dark-adaptation	–56
Low pulse pressure	–36
High PGR initial resistance	34
Little distraction by noise	33
Order of sessions	34
Observer Ratings:	
Persevering, conscientious	37
Accommodating, communicative	24
Self Ratings:	
Persistent, steady, not excited	–33
Kept to one track in conversation	32
Not sensitive, worried or upset	28

slowness, irritability and depression, might then the prominence of the last two be ascribed to some heightening of super ego activity by sub-thyroid conditions, causing conflict with the slowness and diminished capacity to perform? On the other hand, if this should be a fatigue syndrome (as acid saliva could suggest) rather than a sub-thyroid phase, the same kind of hypothesis would need to be invoked as with P. U. I. 4, namely, psychological aspiration causing physiological exhaustion.

A twelfth pattern, worthy of mention, if not of a table, is No. 2 in study C2[20], which is one of carefree emotionality with high fluency and high frequency of laughter, as set out elsewhere[20]. It has the second largest variance in that study, but so far has found no match in any other study.

One must bear in mind that all these factors are in the realm of general personality variables, assessed in behavior rating, questionnaire and objective psychological and physiological test variables. As far as is yet known, both abilities and interests lie outside this realm of measurement. Beyond the five rating, questionnaire and 30 objective personality and physiology test factors condensed in the 12 factors above, seven of which manifest substantial invariance from study to study, there remain for listing eight P– technique patterns found in strictly dynamic variables, i.e., in objectively measured interest attitudes, by computing from function fluctuation observed from day to day.

Actually, ten such factors were found in the pioneer P-technique study, on 20 objectively-measured attitude-interest variables, repeated for 80 occasions over 40 days of varying circumstance[19]. Space precludes setting out the loading patterns, but it may be said that seven of them coincide adequately with R-technique patterns twice found in identical variables[12, 21]. Five of these confirmed patterns are ergs (primary drive patterns), namely, sex, fear, narcism, protective succorance, and self assertion; one appears to be fatigue or a stimulation-avoiding drive, and one a sentiment structure, expressing the self concept. As no physiological associations for these have yet been investigated they will not be discussed further here, but will be listed, for completeness, in the summary of P-technique factors in Table 12 below.

6. DISCUSSION OF RESULTS IN RELATION TO METHODOLOGICAL IMPROVEMENTS

Our primary, restricted objective has been to crystallize in *quantitatively defined patterns* those factors found in P-technique which can be considered reasonably confirmed. These can be starting points for more definitive, intensive, experimental and statistical work on each. Secondarily, we have sought to relate P-technique patterns

where possible to known R-technique patterns. In this we are hampered by the absence of any up-to-date survey of physiological factors, improving on that made in 1946 [11] and paralleling those recently made on personality factors [16, 33].

Any neat integration of this core of P-technique consistencies with results of other approaches is at present difficult. There exist several controlled experiments relating one or two variables at a time; some correlation studies showing certain psycho-physiological clusters; and an accumulation of clinical evidence, with and without controlled medication.

The first two can be confirmed by factor analysis but cannot in themselves return the confirmation, as a correlation can be restored from a factor analysis, but a factor analysis cannot be established from a few correlations. Whenever possible we have tried to integrate these fragmentary indications with ours, but since a single correlation cluster may, for ought we know, hide two distinct factors it is a procedure beset by frustrating doubts. Nevertheless, as indicated, there has been far more consensus than contradiction.

On the other hand, clinical writings on psychological associations of endocrines suffer from uncertainties about defining the physiological variables themselves [57] and as a recent symposium shows [65] are contradictory to the point of absurdity regarding psychological manifestations, except for the main disease syndromes. One therapist regularly cures depressions by cortico-steroid injections, another finds depressive symptoms generally intensified, and so on. Naturally there are many complicating circumstances [55], too numerous to discuss systematically here, but illustratable by such matters as the patient's purely psychological reaction to illness of any kind, the question of whether glandular deficiency really existed before medication, the gross overdoses employed relative to the smaller natural ranges in P-technique, the role of compensatory reactions to outside interference, and the uncertainty as to what functions have been removed in extirpations. But out-topping all other sources of confusion is that to which Beach has drawn attention [3, p. 150] when he says "We are gathered together to talk about the relationship between endocrine functions and behavioral functions, but with few exceptions we seem to be much more sure of ourselves on the endocrinological side than we are on the psychological side . . . Our need is for tests that can be carried out in a standardized fashion . . ." Our results suggest that with suitable personality tests clear patterns can be established. They also indicate that such expressions as "depression", "anxiety", based on different verbal usages, can confuse quite distinct forms, (e.g., our A, F, H and L factors might all (in their negative directions) be called "anxious, withdrawn"). These dimensions can only be separated factor-analytically, and we have good reason to believe that they would behave quite differently in relation to the same endocrine medication.

Just as the psychologist, in these inter-disciplinary studies, is disturbed by the looseness of the estimates of psychological conditions, so the medical clinician in turn is inclined to raise questions about the application of precise factor analysis to physiological measures which he considers may be unreliable. However, on this last point one must bear in mind that the medical clinician has been inclined to equate "significant" with "out of the normal range". A measurement within the normal range loses clinical interest for him, and he is inclined, in dismissing it, to imply that variability in this range is mere error of measurement. The psychologist knows that the question of significance can be examined in terms of analysis of variance or of the consistency coefficient and its comparison with the stability coefficient. The original reliability coefficient tables [22] and score ranges show coefficients with a central tendency of about 0.85 for the physiological measures made by hospital technicians in the studies reviewed here. The normal range in *physiological* measures has therefore no more "error" than the psychologist is accustomed to in measuring the normal range of *psychological* measures. The "function fluctuation" from day to day is of course contaminated with error, as in the psychological measures; but the final proof that it is not *all* error, nor extends to a disabling fraction, is shown by the consistency

of the factor patterns obtainable. Loadings of this magnitude and simple structures of this consistency just could not be built upon measures randomly distributed. The proof of the reality of the patterns is the demonstration that they can be replicated by different experiments. Some failure to appreciate this perhaps arises from lack of statistical awareness that obtaining a whole pattern is a far more improbable event than obtaining a difference in a single measured variable in an expected direction.

Experimental, clinical and factor analytic researches alike, however, admittedly need to operate with more exact measuring devices and particularly more refined endocrine concepts, as soon as biochemical advances make these possible. Pincus points out that the schizophrenic "secretes diminished amounts of cortico-steroid, which differs qualitatively as well as quantitatively from the secretory product of normals' adrenals"[50 p. 163], while Reiss[51] notes different trends, in depression, in different secretions from the same gland. Points of this kind are completely neglected in the prolific theorizing by some writers of psychological textbooks, in which they toss the physiologists' terms around—adrenal function, autonomic, vagotonic, stress —with abandon. Any psychologist who wants to proceed to serious psychosomatic theorizing must first, in Dr. Johnson's well-known phrase "clear his mind of cant", and he may be helped in this attention to more realistic but complex patterns if, as here, he refers to them by some symbol, until the popular associations of terms have extinguished from his thinking.

The clinical suggestion to concentrate on more refined physiological concepts and measurements applies just as strongly to plans for the advance of P-technique factorization techniques. Now that substantial relationships to psychological patterns have been demonstrated in unitary factors, it should be worthwhile to invest in much more extensive and intensive measurement of physiological variables, to check the tentative hypotheses now necessarily made on too bare an array of physiological variables. For example, more (and more diverse) manifestations of thyroxin circulation need to be taken into the definition of Factor 11 above; and, in view of the large role of serum cholinesterase in Factors 1, 2, and 9, and how little is known about its source and action, it would be desirable to include in factorizations any possible physiological variables (e.g., of liver metabolism) that might throw light thereon.

Incidentally, it is not argued here that P-technique is in *all* respects methodologically superior to controlled experiment and uncontrolled clinical observation, though it is unquestionably superior to correlation cluster analysis. However, its only disadvantage relative to the other available methods, that of missing evidence about the causal, temporal sequence can in fact be readily remedied in P-technique by what has been called the "condition-response" design[13]. Similarly the advantage of the magnified manifestations occurring in clinical work can be gained by applying P-technique to clinical cases. The condition-response design would apply the endocrine medication (or other influences) in controlled fashion, while allowing the factor analysis to tease out the functional unities in the responses, or in stimulus and response together, in *a single correlation matrix.*

In further developments in this field it is desirable that P-technique itself be improved both by some explorations of new hunches and by certain devices that are already clearly evident as possible advances. By way of explorations we need more knowledge as to how a large variety of variables, psychological and physiological, behave generally in longitudinal, repeated testing. For example, we need knowledge as to the presence and effects of cycles, range differences, reliabilities of measurement, and score distributions. Already we have some evidence on ranges and frequencies of fluctuation from the early, simple, uncorrelated longitudinal studies stimulated by Flugal and Philpott[49] and especially from the recent multivariate, quantitative, physiological work of Darrow and Kris[43]. The latter have shown, for example, that the variance *within* days is about five times as great as that *between* days, for a fairly typical set of variables, and that both the variances are significant relative to error of measurement. The latter confirms the position we have taken

above *vis-a-vis* medical criticism of the reliability of the normal range of physiological variables and the former throws light on some of the difference of factors between H and K studies here, on the one hand, with their big diurnal range, and the other studies depending more on inter-diurnal variance on the other. Again these studies show that certain formal rhythms [42, 43, 49] exist for particular variables. Their existence suggests (a) that measurement at certain points on these cycles might bring out certain patterns more clearly than would a random set of occasions, as here; (2) that we may need to take special precautions to avoid the Achilles heel of P-technique namely, cycles of equal frequency with certain phase lags which, as Yule pointed out [67] may cause otherwise significant correlations to vanish; and, (3) that we should experiment with time-staggered series in correlation, directed to maximizing all correlations [13] in a matrix, to see if this brings out factor structure more clearly.

Even before these more novel possibilities of improvement are explored much can be done by better implementing the well-known, older needs for adequate technique, most important of which are (1) the obtaining of longitudinal *reliable coefficients* for all variables, by "split-half" devices, giving both consistency and stability coefficients, in a way not yet worked out for many physiological variables, in order to correct resulting factor loadings for attenuation, and, (2) more thorough attention to obtaining good simple structure, with the resulting possibility of proceeding to second-order factors. The present integration shows that simple structure undoubtedly "works" in P-technique, and that it must have priority over researchers' subjective ideas of "meaning". Actually, in the present integration, the usefulness of the consensus of simple structures in several studies was shown in pulling one factor into line, notably in the 17-ketosteroid factor F2 in W. Actually there were several minor adjustments of this kind made in relation to perceived consensus, all of which proved themselves by producing simultaneous improvement in other ways. Because of rotation difficulties it may be doubted whether any single factorization (in any area of science), however good in sample and test conditions, can be finally accepted until checked with a second study, and an independent rotation.

Although the seven pioneer studies here give us seven or eight factor patterns which are in this respect dependable, they need improved definition, partly through attenuation correction, as just indicated, partly through determining more definitely the usual degrees of obliqueness existing between certain factors, but mainly through the inclusion of more variables, particularly variables strategically chosen in relation to the present hypothesis, which may yield high loadings on the factors and permit the factors as such to be measured with adequate validity in clinical and general experiment. Especially is it important, for our understanding of the autonomic system, to insert the critical variables, e.g. pupillary dilation, to test our hypothesis that Wenger's two factors are really four.

Though such broad P-technique studies extending the present variables to permit a fresh grip on the factor are vitally necessary, we may already claim to have reached the stage in the seven general and eight motivation P-technique factors, as in the 12 invariant R-technique factors [16], at which experimenters can confidently take one factor and study its variables and structure intensively. This needs to be done factor analytically, seeking for more saturated variables, and by controlled experiment. Referring to particular factors, one would particularly suggest studies: (1) More clearly to separate the two diurnal trend factors—P. U. I. 2, warming up, and P. U. I. 3, fatigue; (2) more definitely to separate the four autonomic factors in terms of neurological and endocrine studies; (3) with accurately estimated and applied physiological variables, notably ACTH, DCA and STH, to clarify the nature and sequence of the "emotional control-high corticosteroid" reaction represented by P. U. I. 4; and (4) on anxiety, in any of the contexts currently interesting researchers, but with measurement of the distinct factors in anxiety level which P-technique indicates, namely, anxiety-depression, P. U. I. 5 (as sympathetic overactivity), P. U. I 8 as the "general neuroticism" factor, and P. U. I. 11 as a possible thyroid reaction.

7. Expectations of Factor Invariance Over R- and P-technique Experiments

A surprising outcome of the present integration is the high degree of apparent matching of P-technique factors with R-technique rating and objective test factors. For in the latter realm itself, the alignment of rating (or life record criterion) factors and test factors has been quite poor[12]. Incidentally, since so few physiological associates have yet been explored for R-technique temperament factors, the present physiology-rich patterns, as matched with R-technique, may help psychosomatic research in the latter realm by indicating the physiological associates that are "best bets" for initial investigation there. However, it must be stressed that factor matching both within P-technique and across to R-technique is a tricky matter, involving wide statistical considerations. Satisfactory decisions can be made only when experiment is re-centered with what the present survey now shows to be the vital variables for these factors, and, additionally, with dual or split-half measures from which reliability coefficients can be calculated, to determine how much of the difference of loading from expected values could be due to unreliability of measurement.

The use of the term "invariant factors" in the title of this article must be understood in relation to the above considerations. The aim of factor integration is to bring evidence of replicatable, invariant factor loading patterns; but the actual experimental patterns here set out are not precisely identical, nor is it required that they should be. Invariance is a statement of hypothesis: that the degree of departure from identity is in these cases only such as would be expected from sampling and experimental error, and—especially in matches from R– to P— by such additional systematic effects as are discussed below. Actually, the requisite degree of invariance to justify the hypothesis that the same, identical influence is behind the patterns in the different experiments is claimed here only within the P-technique matchings, and for that reason the precaution has been adopted of separate Universal Index numbers for P– and R–technique series.

When all statistical improvements have been made to permit allowing for random error in recognizing those functional unities which show in both "within" and "between" individual variation, it must be realized that we should still expect certain *systematic* differences of loading pattern between R- and P-technique factor expressions of any influence. However, by proper consideration of the total situation it should be possible to recognize how these would affect a factor match. The chief considerations are: (1) That factors identified as physiological or ergic patterns should have relatively large variance in P-technique, whereas those due to abilities, complex personality structure and learned sentiment patterns may be barely reproducible because of slightness of intra-individual fluctuation. This theoretical expectation is supported by the present survey where temperamental patterns like H, F and A are large, C and G are small, and ability or learning patterns like B and K[11, 12] are absent, and where seven ergic but only one sentiment pattern are apparent. It is supported also in that three objective test patterns U. I. 1, 20, and 28 (intelligence and two culturally acquired patterns) do not appear in P-technique. (2) The prominence of factors should relate intelligibly to the kind of sampling of occasions made. A large within-day range may magnify fatigue factors, and, at the opposite extreme, occasions sampled over months should magnify a seasonal trend and may even bring out developmental factors. (3) Factors which show very great P-technique variance, especially diurnally, should not be expected to involve permanent personality characteristics. (4) Some variables will show in a factor pattern perceived in one technique but not in the other because they may arise, in R-technique, only from *prolonged* action of the factor influence concerned upon these variables. For example, size of vocabulary strongly loads the R-technique intelligence factor, presumably through long-term learning, but momentary fluctuations in intelligence would not be expected to influence it so much. The P-technique factor should thus distinguish, for example, the patterns and concepts of fluid from crystallized intelligence[10], presenting us, incidentally with an additional operation to augment the meaning of this distinction.

Fifthly and lastly, one must never forget that the P-technique factor strictly belongs to a different species from the R-technique factor, in that it is not a common but a unique trait. Indeed it is this which makes the method of particular value to the clinician in unravelling *individual* dynamic fixations[14]. However, in terms of matching, we may expect that the subject's individuality, as a departure from the R-technique pattern, would, according to general psychological knowledge, show itself far more in *dynamic* than in general personality or physiological factors. In any realm, however, we should expect the R-technique patterns to be the central tendency of a sample of P-technique patterns. This raises the vexing question of how much scientific generalization may be attempted from two or three P-technique cases, compared with, say two or three hundred R-technique cases, but involving the same number of measurements. (As would be the case, say, if each P-technique were based on 100 occasions). It has just been seen that, in fact, good general agreement on R- and P-technique patterns can be perceived, to the satisfaction of common sense, when based on three or four P-technique cases. One may nevertheless ask for statistical means of deciding how much of the difference is real and how much mere error, and just what degree of generalizability can be given to the common pattern.

It is interesting that the emphasis on large samples has grown up largely among psychologists, whereas the medical man, and, still more, the chemist, is content to believe that the process which happens in one case will always happen. The great anatomists did not need to dissect 100 cadavers to generalize about, say, the form and function of the pulmonary artery in man, and medicine based a whole physiology of the stomach on Beaumont's observations on one case. The rationale for these apparently diverse standards is that some general and vague previous experience of the standard deviation existing in different phenomena is implied in such judgments and that psychological individuation is accepted to be greater than physiological individuation. Within psychology it would seem that about learning patterns derived from individual life history, as seen by the psychotherapist, we cannot safely generalize, but that to the extent to which we deal with unlearned patterns, or with reflections of cultural institutions, or, more widely, with *relations among* elements, we are dealing with what is more likely to be common and true of all men. Blurring by superposing many different individual patterns may be as much responsible as small variance for the difficulty of getting acquired, common patterns in R-technique which have sometimes been recognized as repeated unique patterns in P-technique. Repetition of *identical P-technique* variable studies on several individuals is now much to be desired if we are to determine the relative roles of error and true individuality in these differences. Meanwhile, it is evident from the similarity of only two or three cases that at least the above seven more invariant psychosomatic patterns and the six ergic and one metanergic patterns are sufficiently rooted in general human physiology, constitution and culture to be studied, if necessary, intensively in the single case.

Summary

(1) From five P-technique studies with psychological and physiological variables, each rotated independently (initially), seven factor patterns have been found which are recognizably repeated, i.e., reasonably invariant from study to study, and five others which are sufficiently statistically substantial to justify some statement and hypothesis formation.

(2) These factors, and seven others from a sixth study confined to motivation measurements, are for the most part matchable with major factors found in R-technique studies, in ratings, physiological variables, objective personality tests and motivation measures. They also agree, as to parts of their total pattern, with the more fragmentary connections previously found in pairs of correlated variables, in controlled univariate experiment, and particularly, with triads or more of physiological "syndromes" found in basic physiological research.

(3) The patterns thus examined have been given P-technique "Universal Index" symbols as listed in table 12.

TABLE 12. SUMMARY OF P-TECHNIQUE PATTERNS.

A. General Psychosomatic Factors

Proposed P-technique Universal Index No.	Physiological Title	Behavior Rating Factor name and Symbol	Matched R-technique Objective Test Factor Universal Index No.
P. U. I. 1	Parasympathetic Resilience	Adventurous Cyclothymia (H)	(U. I. -17, 21)
P. U. I. 2	Low Cholinesterase	Desurgency vs. Surgency (F-)	U. I. -22
P. U. I. 3	G.A.S. Stage of Resistance	General Diurnal Fatigue	———
P. U. I. 4	G.A.S. Alarm Reaction	Control (Ego Strength) (C)	U. I. 26
P. U. I. 5	Adrenergic Reactivity Level	Cyclothyme State (A)	U. I. -19
P. U. I. 6	Warm Season Trend	Warm Season Trend, plus Practice	———
P. U. I. 7	Basophil-Neutrophil	———	———
P. F. 8	———	Untapped Reserves vs. Overwroughtness	U. I. 23
P. F. 9	Autonomic Inactivity	Low Metabolism Paranoid Trend (L)	———
P. F. 10	Autonomic Reflex	Nervous Disposition	U. I. 29
P. F. 11	Hypothyroid	Super Ego (G)	(U. I. 17, -21)
P. F. 12	———	"Carefree Emotionality"	———

B. Motivational, Ergic and Sentiment Factors

P. U. I. 8	Sex erg,	(Need level)	Matched in R-technique			
P. U. I. 9	Fear-escape erg,	"	"	"	"	"
P. U. I. 10	Narcistic erg,	"	"	"	"	"
P. U. I. 11	Protective, succorant erg,	"	"	"	"	"
P. U. I. 12	Self assertive erg,	"	"	"	"	"
P. U. I. 13	Stimulation avoiding erg,	"	"	"	"	"
P. U. I. 14	Self-sentiment structure	"	"	"	"	"

(4) The first set of factors indicate substantial psychosomatic connections and would seem to constitute the entities needed for measurement in experimental studies of anxiety, stress, autonomic influences, etc. Some statistical considerations governing matching, invariance of form, interpretation and generalization in connection with such patterns are set out.

(5) While the main patterns agree with tolerably well-known physiological functional concepts, others are as new to the physiologist as to the psychologist. Space denies us speculation about the interpretation of these, but this does not interfere with our primary objective which has been the setting out of a number of relatively invariant factor loading patterns, as a definite basis for experimental and general research, and for the termination of the confusion that has existed among these similar patterns. In effect we finish with twelve hypotheses, each stating that a particular psycho-physiological pattern (more extensive than any present single P-technique research would indicate) will be found in further P-technique and clinical research, as a necessary pivot for any functional studies. It is assumed that each has arisen through "function fluctuation" of a functional unity, observable, or already observed in many cases, also in the factor analytic study of individual differences.

(6) Since growth and integration of this area of knowledge depend on use of exactly similar "marker" tests, and since repeatable tests are not easily devised or selected, the present writer has set up a P-technique battery of the better tests for the principal personality and motivational factors, of which some investigators have already availed themselves. However, there is also need for development of about 100 *equivalently* (*group*) *standardized* tests for any one variable, for the situation when one test does not bear repetition without invalidation, e.g., a hundred intelligence tests of known score equivalence, or a hundred rigidity tests standardized on the same large population. This is much to ask, but until such tests of known equivalence are available some important R-technique areas will not become available to P-technique checks.

Bibliography

1. Alexander, F. *Psychosomatic medicine.* New York: Norton, 1950.
2. Baldwin, A. L. The study of individual personality by means of intra-individual correlation. *J. Personal.*, 14, 1946, 151-169.
3. Beach, F. A. Discussion on personality in Addison's disease. p. 150 in *Ciba Foundation Colloquia on Endocrinology.* Edit: Wolstenholme, G. E. W., New York: Blakeston, 1952.
4. Best, C. H., and Taylor, N. B. *The physiological basis of medical practice.* 4th Edition. Baltimore: Williams and Wilkins, 1945.
5. Bills, A. G. *The psychology of efficiency.* New York: Harper, 1943.
6. Broun, G. O. Blood destruction during exercise. I. Blood changes occurring in the course of a single day of exercise. *J. Exp. Med.*, 36, 1922, 481-500.
7. Browne, J. S. L. The interplay between endocrine disturbances and psychological aberrations. In Ciba Foundation Colloquia on Endocrinology. Edit: Wolstenholme, G. E. W., New York: Blakeston, 1952, 128-140.
8. Brozek, J., and Keys, A. Flicker fusion as a test of fatigue. *J. Indust. Hygiene*, 26, 1944, 169-174.
9. Cattell, R. B. The significance of the actual resistances in psychogalvanic experiments. *Brit. J. Psychol.*, 19, 1928, 34-43.
10. Cattell, R. B. The measurement of adult intelligence. *Psychol. Bull.*, 40, 1943, 153-193.
11. Cattell, R. B. *The description and measurement of personality.* New York: World Book Company, 1946.
12. Cattell, R. B. *Personality: A systematic theoretical and factual study.* New York: McGraw-Hill, 1950.
13. Cattell, R. B. *Factor analysis.* New York: Harper, 1952.
14. Cattell, R. B. P-technique factorization and the determination of individual dynamic structure. *J. clin. Psychol.*, 8, 1952, 5-10.
15. Cattell, R. B. A universal index for psychological factors. Adv. Publ. No. 3, Lab. Person. Assess., University of Illinois, Urbana, 1953.
16. Cattell, R. B. The principal replicated factors discovered in objective tests. *J. abnorm. and soc. Psychol.*, 1955, 50, 291-314.
17. Cattell, R. B., Blewett, D. B., and Beloff, J. The inheritance of personality: A multiple variance analysis determination of approximate nature-nurture ratios for primary personality factors in Q-data. (In press).
18. Cattell, R. B., Cattell, A. K. S., and Rhymer, R. M. P-technique demonstrated in determining psycho-physiological source traits in a normal individual. *Psychometrika*, 12, 1947, 267-288.
19. Cattell, R. B., and Cross, K. P. Comparison of the ergic and self-sentiment structures found in dynamic traits by R- and P-techniques. *J. Personal.*, 21, 1952, 250-271.
20. Cattell, R. B., and Luborsky, L. B. P-technique demonstrated as a new clinical method for determining personality and symptom structure. *J. genet. psychol.*, 42, 1950, 3-24.
21. Cattell, R. B., and Miller, A. A confirmation of the ergic and self-sentiment patterns among dynamic traits (attitude variables) by R-technique. *Brit. J. Psychol.*, 43, 1952, 280-294.
22. Cattell, R. B., and Williams, H. V. M. P-technique, a new statistical device for analyzing functional unities in the intact organism. *Brit. J. Preventive and Soc. Med.*, 7, 1953, 141-153.
23. Dalton, A. J., and Selye, H. The blood picture during the alarm reaction. Folia Haemat. Lpy., 62, 1939, 397-407.
24. Darling, R. P. Autonomic action in relation to personality traits of children. *J. abnorm. and soc. Psychol.*, 35, 1940, 246-260.
25. Darrow, C. W., and Heath, L. L. Reaction tendencies relating to personality. In Lashley, R. S. *Studies in the dynamics of behavior.* Chicago: Univ. Chicago Press, 1932.
26. Dodge, R. M. Elementary conditions of human variability. E. K. Adams Fund for Physical Research. No. 10. Columbia University Press, 1927.
27. Dunbar, H. R. *The emotions and bodily changes.* New York: Columbia Univer. Press, 1938.
28. Egoroff, A. Z. Die veranderung des blutbildes wahrend der muskelarbeit bei gesunden (die myogene leukocytose). *Klin. Med.*, 100, 1924, 485.
29. Eysenck, H. J. *Dimensions of personality.* London: Kegan Paul, 1947.
30. Forbes, A. P., Donaldson, E. C., Riefenstein, E. C., and Albright, F. The effect of trauma and disease on the 17-ketosteroid excretion in man. *J. clin. Endocr.*, 7, 1947, 264-288.
31. Freeman, G. L., and Katzoff, E. T. Individual differences in physiological reaction to stimulation and their relation to other measures of emotionality. *J. exp. Psychol.*, 31, 1942, 527-537.
32. Freeman, G. L., and Elmadjian, F. The relations between blood sugar and lymphocyte levels in normal and psychotic subjects. *Psychosom. Med.*, 9, 1947, 226-233.
33. French, J. W. The description of personality measurements in terms of rotated factors. Princeton: Educational Testing Service, 1953.
34. Fulton, J. T. *Physiology of the nervous system.* New York: Oxford Press, 1938.
35. Glaser, F. Ueber den kinischen nachweis psychophysicher reaktionen. Die appetitsaft leukozytose. *Med. Klinik*, 20, 1924, 535-537.
36. Haverland, E. M. An experimental analysis by P-technique of some functionally unitary varieties of fatigue. M. A. Thesis. Urbana, Illinois: University of Illinois, 1954.
37. Herrington, L. P. The relation of physiological and social indices of activity level. In McNemar, Q, and Merrill, M. A. *Studies in personality.* Stanford University Press, 1942.

38. HEUSNER, W. W. Changes in physical fitness variables associated with intensive training in swimming. Ph.D. Thesis. University of Illinois.
39. HOAGLAND, H. Stress and the adrenal cortex, with special reference to potassium metabolism. *Psychosom. Med.*, 12, 1950, 142-148.
40. HOLMES, T. H., et al. The nose: An experimental study of reactions within the nose in human subjects during varying life experiences. C. C. Thomas, 1950.
41. IZQUIERDO, J. J., and CANNON, W. B. Emotional polycythemia in relation to sympathetic and medulliadrenal action on the spleen. *Amer. J. Physiol.*, 84, 1928, 545-562.
42. KARVONEN, M. J., and KUNNAS, M. Factor analysis of haematological changes in heavy manual work. *Acta Physiologica Scandinavica*, 29, 1953, 220-231.
43. KRIS, C. Diurnal fluctuations in body temperature, heartrate, and work output. M. A. Thesis. Illinois Institute of Technology, 1953.
44. KUNTZ, A. *The autonomic nervous system.* 3rd Edition. Philadelphia: Lea and Febiger, 1945.
45. MEELAND, T. An investigation of hypotheses for distinguishing personality factors A, F and H. Ph.D. Thesis, University of Illinois, 1952.
46. MILLER, E. B. Physical and physiological studies, pp. 25-87 in Sanford, R. N. and others, *Physique, Personality and Scholarship.*
47. MOHR, F. *Psychophysische Behandlungsmethode.* Leipzig: Hirzel, 1925.
48. MORGAN, C. T., and STELLAR, E. *Physiological Psychology.* New York: McGraw-Hill, 1950.
49. PHILPOTT, S. J. F. Fluctuations in human output. *Brit. J. Psychol.*, Monogr. Suppl. No. 17, London, 1933.
50. PINCUS, G. Adrenal function in schizophrenic men. Ciba Foundation Colloquia on Endocrinology. Edit: Walstenholme, G. E. W., 1954-167. New York: Blakeston, 1952.
51. REISS, M. Psychological changes connected with spontaneous and experimentally produced alternatives in steroid hormone metabolism. In Ciba Foundation Colloquia on Endocrinology. Edit: Walstenholme, G. E. W., 128-140. New York: Blakeston, 1952.
52. RICHTER, D., and LEE, M. Serum cholinesterase and anxiety. *J. ment. Sci.* 88, 1942, 435-439.
53. SARGENT, F. A critique of homeostasis: Season and metabolism. *Archiv. f. Meteor., Geophys. u. Bioklin.*, 3, 1951, 289-296.
54. SELYE, H. *The story of the adaptation syndrome.* Montreal: Acta, Inc., 1952.
55. SPEIRS, R. S., and SULLIVAN, J. Comparison of epinephrin and ACTH action on the circulating eosinophils of mice. *Amer. J. Physiol.*, 173, 1953, 282-288.
56. STRONGIN, E. T. A laboratory method of diagnosing manic-depressive depression. *Psychol. Bull.*, 39, 1942, 509.
7. SUMNER, J. B., and MYRBACK, K. *The enzymes.* New York: Academy Press, 1950.
58. TOMASSON, H. Psychische beeinflussung des serum calcium spiegels. *Klin. Wschr.* 3, 1924, 2055-2056.
59. TOD, H., and JONES, M. S. A study of the cholinesterase activity in nervous and mental disorders. *Quart. J. Med.*, 6, 1937, 1-3.
60. TOD, J. C., and SANFORD, A. M. *Clinical diagnosis by laboratory methods.* Philadelphia: Saunders, 1939.
61. VORHAUS, L. J., and KARK, R. M. Serum cholinesterase in health and disease. *Amer. J. Med.*, 14, 1953, 707-719.
62. WENGER, M. A. Interrelations among some physiological variables. *Psychol. Bull.*, 37, 1940, 466.
63. WENGER, M. A. Studies of autonomic balance in Army Air Force Personnel. *Comp. Psychol. Monogr.*, 1948, 101, 1-111.
64. WILLIAMS, H. V. A determination of psychosomatic functional unities in personality by means of P-technique. *J. soc. Psychol.*, 39, 1954.
65. WOLSTENHOLME, G. E. W., Ciba Foundation Colloquia on Endocrinology, Vol. 3, Philadelphia: Blakeston, 1952.
66. WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.
67. YULE, C. U. Why we sometimes get nonsense correlations between time series. *J. Royal Statist. Association*, 89, 1926, 1-64.

AN INVESTIGATION OF AN ABNORMALITY OF COGNITIVE FUNCTION IN A COOPERATIVE YOUNG PSYCHOTIC: AN EXAMPLE OF THE APPLICATION OF EXPERIMENTAL METHOD TO THE SINGLE CASE

M. B. SHAPIRO AND E. H. NELSON

Institute of Psychiatry, Maudsley Hospital, London

The Single Case In Experimental Psychiatry

It has been argued elsewhere[8, 9, 11] that the systematic objective investigation of a single abnormality in a single case should lead to findings which are relevant to the general function of the individual concerned and which can be generalised to many other individuals. The first attempt to check this viewpoint is reported in the papers quoted above. The investigation reported in these papers began with the experimental investigation of an apparent anomaly of perceptual function in a psychiatric patient. As a result, some of the laws determining the appearance of the anomaly were established. Further investigation showed that the anomaly appeared most frequently amongst brain damaged patients. An explanation was then formulated, accounting for the facts that the anomaly obeyed these particular laws and appeared most frequently in brain damaged patients. This explanation, which had implications for the general adjustment of brain damaged patients, also permitted precise deductions. One of these deductions was the conditions which would make a normal person also show the perceptual anomaly and do so according to the same laws. Experiment confirmed these deductions.

It might be argued that the apparently successful outcome of this investigation was partly due to chance and partly due to the fact that we were dealing with an anomaly which was related to brain damage. It was, therefore, necessary to start another investigation of this type in a patient who was suffering from a functional disorder. This paper reports such an investigation.

Before reporting this investigation it is necessary to deal briefly with the underlying theoretical conceptions. The first of these is that the individual's total behavior and experience is a function of the interaction of various aspects of psycho-physiological function. The experimental investigation of the causes of any single abnormality of psychological function may, therefore, lead to an improved understanding of the patient's abnormality as a whole. The second conception is that the understanding of the psychological dysfunction of any single patient may be relevant to the understanding of the disorders of many other patients. It should be stated at this point that any function is regarded as psychological if the subject reports that conscious experience has been in any way involved.

These conceptions have to be made explicit because many psychologists and psychiatrists appear to work on the unproven assumption that the precise investigation of any single aspect of function in a single case *must* result in unique findings, findings which will be incapable of generalization, either to other aspects of the individual's function, or to the same function in other people. This assumption has done much to paralyse effective experimental work in clinical psychology.

The effectiveness of an investigation of a single aspect of function in a single case depends, however, on the rigour with which it is carried out. To quote from a previous paper[9]: "By rigour is meant, first of all, that no experiment is ever conducted unless a formulated explanatory hypothesis is being tested. Secondly, no hypothesis is ever considered unless testable deductions can be made from it. Such deductions must result in experiments in which the phenomenon concerned varies its appearance in a manner which confirms or contradicts the hypothesis. Such an approach would, of course, become an empty drill, if the hypotheses which are developed are not explanatory and capable of wider application beyond the immediate limits of the point of the investigation."

The Initial Observation

The initial observation which started the investigation reported in this paper was the fact that a 22 year old male psychiatric patient, who had obtained well above average scores on cognitive tests such as the Wechsler, was unable to understand simple material when he read it aloud. The Wechsler-Bellevue scores are given below:

Full Scale I. Q. 118

Verbal Scale I. Q. 120		Performance Scale I. Q. 113	
Comprehension	13	Picture Arrangement	11
Digit Span	11	Picture Completion	11
Arithmetic	10	Block Design	15
Similarities	13	Object Assembly	14
Vocabulary	15	Digit Symbol	9

Psychiatric Summary of the Subject

The patient had his first breakdown at the age of 20 while he was studying for his degree as a resident student of a university. This breakdown was preceded by a period of inability to concentrate on his studies, and of apparent "withdrawal" from his fellow students, teachers and relatives. Under psychiatric observation he was, in addition, reported to have rapid variations in mood, "significant thoughts", suspiciousness and panics. He was given 40 insulin comas, and made an apparent recovery. After spending some months hitch-hiking and working his way around a country abroad he returned to his university, and recommenced his studies in the autumn term. He felt very well until December of that year when he again found difficulty with his concentration, with attendant anxiety about his mental state. He finally became "depressed" and "perplexed" and was given another 40 insulin comas and 6 E. C. T. with considerable improvement. It was at this phase that he was seen by our psychological department. In the testing situation he appeared to be an extremely cooperative, pleasant, sensible and persistent person.

He is reported to have been a happy child but shy, easily frightened and crying easily. No other abnormal traits are reported. At school he was a good average scholar, active socially and good at games. He passed school certificate but failed higher schools. On leaving school he went into the Army for two years and obtained a commission. His father is a successful professional worker, and the home environment is apparently kindly and cultured. He is reported to have shown no interest in sex and not to have had either homo- or heterosexual experiences. The patient had complained that he found it difficult to read books and in order to check this complaint the reading comprehension tests were given.

Objective Confirmation of the Complaint

In order to confirm the patient's complaint, two tests were given: the Schonell Prose Reading Test "My Dog" and the Schonell Silent Reading Test A[7].

Schonell Prose Reading Test "My Dog"[7]. In this test the subject has to read aloud a simple story. Three scores are recorded: the number of mistakes in actual reading, the time taken to read the passage, and the number of standard questions on the content of the story which can be answered correctly.

Results: (a) *Reading Mistakes.* <9.5 years (the ceiling of this test is at 9.5 years).

(b) *Reading Speed.* Between 8.5 years and 9.5 years (the ceiling is at 9.5 years).

(c) *Answering Oral Questions on the Test.* Four out of 15 were answered correctly, giving a reading age of less than 6 years.

It seemed from these results that the patient could make correct oral responses to print, but did so slowly and with almost no comprehension of what he had read.

In the light of his full scale Wechsler score and his school record, this lack of comprehension was regarded as being abnormal and as confirming the patient's claim that he had only been able to read through one book in the whole of his life. It was decided to make the apparent abnormality of performance on this reading test the starting point of the investigation.

Schonell Silent Reading Test A [7]. In this test the subject is left to read a number of passages to himself. Each passage is followed by printed questions which the patient has to answer in writing. The correctness of the subject's answers to these questions gives a measure of the subject's comprehension of what he has read. Two scores are derived: accuracy and speed. On the first he broke the ceiling of the test which goes up to 13½ year level. On the second he obtained a score between 10 and 11 years.

Preliminary Hypotheses

Seven explanatory hypotheses were set up in the early stages of the investigation. For the sake of brevity of exposition they will be stated at the outset, though in fact they were formulated at different times. It should be noted that these hypotheses were not mutually exclusive. The hypotheses were:

(i) The patient had a specific reading defect.
(ii) The patient fatigued very quickly.
(iii) He had difficulty in attaching meaning to words.
(iv) The patient had a reading comprehension difficulty which was characteristic of some schizophrenics.
(v) The patient was extremely slow in reading comprehension.
(vi) "My Dog" is psychologically an inappropriate test for adults. It could be argued that reading simple material aloud was both an unaccustomed and embarrassing task for adults.

Testing the Hypotheses of a Specific Reading Defect and Fatiguability

A number of actions were taken to investigate these two hypotheses:

1. The history of the patient was investigated. This revealed that the patient had never been backward in reading. He claimed, however, that he had had as a child a tendency to read from right to left, to make reversals such as "on" for "no". His cultured family had teased him, he said, for his so-called backwardness.
2. His present reading habits were checked. No clear-cut abnormalities could be observed such as eye-voice lag, or reading words singly. The only thing which was clearly confirmed was his claimed inability to take in what he read.
3. In order to establish the limit of his comprehension in silent reading as against reading aloud, he was given the Schonell Silent Reading Test B. This test has a ceiling of 15 years.

At the same time the hypothesis of easy fatiguability was checked by preceding the Reading test with the first two subtests of the Wechsler Verbal Scale and then following the reading test immediately with the rest of the Wechsler test. The Wechsler test results were required anyway, as it was necessary to check his present intellectual level. The patient appeared to be as efficient in the latter part of the Wechsler as in the early part. Grossly abnormal fatiguability was not observable here.

On the Silent Reading test he obtained a reading age of 12 years, answering 28 out of the 42 questions in the test correctly. He did not finish the test in the time limit of 15 minutes.

4. To ensure that there was not some actual disability which prevented the patient from finishing the Schonell Silent Reading Test B, and to check the hypothesis of slowness of reading comprehension, the patient was given the test to complete in his own time on the following day. This he did, and he obtained the maximum reading age of 15 years 2 months on the test. For the purposes of this investigation the hypothesis of a specific reading defect was dropped therefore at this stage.

Checking the Inappropriateness of "My Dog" As a Test for Adults

Here again a number of things were done.

1. The patient was given Burt's "King of the Golden River"[1]. This test is similar to "My Dog" in that it consists of relatively simple material which is read aloud and the subject is subsequently tested for comprehension by questions. This test was given to see if there was anything peculiar to the content or the conditions of the administration of "My Dog". The patient in fact obtained a reading age of 7 to 8 years on Burt's test, which was again far below what one would expect from the patient's intelligence level and record.

2. In order to check the standardization and the psychological appropriateness of the reading tests used so far, three of them were given to a sample of normal controls who were like the patient in social status and intellectual level. Most of them were Psychology graduates working in the department. The results were:

(i) *My Dog.* Nine subjects were tested. The range of scores was from 6 to 15 out of 15 questions answered correctly. The mean reading age was 8.5 - 9 years. Some of the controls obtained scores nearly as bad as that of our subject. The raw data have unfortunately been lost.

(ii) *King of the Golden River.*

Frequency	*Reading Age*	*No. Correct Answers*
5	9-10 years	9-11
1	11-12	12-13
1	12-13	13-14
2	13-14	14-16
Patient	7-8	5

Five of our controls obtained a score well below the ceiling of the test and only slightly superior to that of the patient. Only two of the controls just reached the ceiling of the test.

(iii) *Schonell Form B* (*Silent Reading Comprehension*). All nine subjects obtained the maximum score of 15 years 2 months for men and of 15 years for women within the time limit of 15 minutes.

Two conclusions emerge from these results. The first is that the controls found no difficulty with the silent reading situation as measured by Schonell Form B. Thus our patient's slowness of reading comprehension is confirmed.

The second conclusion is that our adult controls found it relatively difficult to comprehend simple material when they themselves read it aloud and that therefore the patient's difficulty in this test is not definitely abnormal. On the other hand we are not in a position to say that the patient showed no abnormality in this test for his score is still worse than any of the controls and the number of controls is too small to decide on the significance of this observation.

The Relation of Comprehension While Reading Aloud To Schizophrenia

To check this relation "My Dog" was given to two schizophrenic patients of well above average intelligence, whose diagnoses were unequivocal.

Diagnosis	*Reading Age*	*Time Taken*	*No. Correct*	*C. A.*
Hebephrenic	6	50″	6	19
Paranoid schiz.	6	60″	6	24
Our patient	less than 6	86″ approx.	4	22

These results lie between our patient's and that of the control groups. Again, we have insufficient data to indicate the degree of abnormality of these findings.

Recapitulation and Reformulation

Four main findings emerged from the data so far gathered:

1. There was a contraindication of the existence of a specific reading defect in the patient.

2. There was some evidence that the patient was not abnormally prone to fatigue in cognitive function.

3. There were good grounds for believing that the patient was relatively slow in the comprehension of printed matter.

4. There was clear evidence that the comprehension of simple material whilst reading it aloud is difficult for some normal adults. (At least they cannot do it better than children.)

The last of these results appeared to rob us of the apparently dramatic and clear cut abnormality which we seemed to have at the outset of the investigation. The disappearance of this abnormality was not, however, complete. Our patient was consistently worse than the controls in the reading tests, and the schizophrenics in turn were worse than the normals on "My Dog", though doing better than our patient. Our numbers were too small and our sampling too restricted to make an accurate estimate of the degree of abnormality of this finding. We were therefore not in a position to decide with any confidence for or against the existence in our patient of an abnormality, in the strict sense of the term, of reading comprehension as measured by "My Dog".

Two possible alternatives emerged from the situation: either to collect a considerable amount of control data on "My Dog" which would give us a reliable measure of the patient's degree of abnormality in this task, or to put forward an explanation of all the available results which could be tested by an adequately standardized test. It was decided to do the second, if possible, as it was the less time consuming procedure.

A General Explanation and Its Validation

Further consideration of the results obtained suggested an explanation. This was that our patient was abnormally slow in any kind of cognitive function. This was indicated by his relative slowness in the silent reading comprehension test, a function which is known to be quite highly correlated with other types of cognitive tests. Our normal controls all exceeded the maximum score on this test. Slowness of cognition could also account for the patient's difficulty in comprehending what he had read aloud. If the patient's rate of responding verbally to print were relatively fast (as it was) then his rate of comprehension could not keep pace with rate of reading aloud. Fortunately an appropriately standardized and validated test was available to check the theory of the patient's cognitive slowness[3, 4]. In fact its existence brought the hypothesis readily to mind. This test was Furneaux's speed and level test of inductive reasoning, which is based on Thurstone's letter series test.

Each item in this test is made up of a series of letters. Each series has a definite pattern of relationships. The subject has to work out the pattern and then complete the series. A simple example would be the AXAYBXBYCXC. The subject has to fill in the one letter indicated by the dot.

In the level test the subject is instructed to take as long with each item as he desires and only to respond with the answer when fully satisfied that it is correct. He is then given six practice items. When this practice is completed the subject goes on to the main part of the level test. This consists of three independent series of items. Each series consists of 7 items, each of which is on a separate card. The items of each series are graded in difficulty level from that passed by 80% of the normal population to that passed by 5%, and the graduation in each series is of the same order. Thus each individual has three opportunities to show the level he can reach.

The speed test is always given after the level test so that differences in learning and practice effects can be reduced. All the items of the speed test can be passed by 80% of the normal population. The subject is required to work as quickly as possible, to answer each item in turn without omitting any items. To facilitate scoring he is asked to make a circle around the last item completed whenever the examiner taps his pencil. This is done every 20 seconds. The essential virtue of the construction of the speed test lies in the fact that only the items which a subject does correctly are scored. The speed test therefore gives a measure of a subject's speed of successful function, and is not complicated by the time he spent on incorrectly completed items.

The results of the test were quite clear. The patient obtained a score of +.57 sigma on the level test. The standardization data were obtained from University undergraduates. On the speed test he obtained a score of −1.5 sigma. The discrepancy of 2.07 sigma between speed and level was extremely unusual. In Furneaux's control data there are 172 students who obtained a score between +.5 and +.6 sigma (this information was given in a private communication and is about to be published). Of these only three obtained a discrepancy of two sigma or more. The hypothesis that our patient was abnormally slow in cognitive function was confirmed.

To check on whether this abnormal discrepancy was a general function of mental disorder, the speed and level tests were given to four diagnosed schizpohrenics:

Patient	*Level**	*Speed**	*Level-Speed Discrepancy**
1	+0.86	−1.66	−2.52
2	+0.86	−3.1	−3.96
3	+1.85	+2.0	+0.15
4	+0.16	−2.56	−3.72

*All results are given in terms of standard deviations of the normal student sample.

Of these patients only No. 3, the one without a discrepancy, was rediagnosed before discharge. The new diagnosis was that of anxiety neurosis. According to Furneaux's data the remaining three discrepancies were at least as abnormal as that given by our patient. We thus appear to be dealing with an abnormality characteristic of at least some schizophrenics.

To check the link of this finding with slowness of reading, one more test was given. The patient was asked to read a story in a magazine. Fifteen questions were prepared to test the subject's comprehension. Our patient took 10½ minutes to read the story, silently to himself, and answered 14 out of 15 questions correctly. Eight normal controls ranged from 2½ to 7 minutes in reading the story and none answered fewer than 14 out of 15 questions correctly. All subjects were asked to read in their own time. It would seem, therefore, that the patient suffered from a general slowness of intellectual function and the reading difficulty may well be related to this slowness.

Testing the Generality of the Finding

It now remained for us to confirm the generality of the finding that some schizophrenics, while suffering little or no impairment of level of cognitive function, would be abnormally slow in carrying out such function. This investigation was carried out by Ogilvie[6].

He gave the Nufferno level and speed tests to 22 patients. The selection of patients was governed by the aim of making them as like our original patient as possible. Their average age was 24 years with a range from 18 to 36 years. They were all considered to be of above average intelligence in the opinion of their psychiatrists and at the time of testing were considered to be schizophrenic. They were all fully cooperative in the testing situation. The short form of the Nufferno Level Test and the interrupted form of the speed test described above were used. The results were compared, as before, with the data previously gathered from the university students.

The results were fully in line with expectation. The mean time per item successfully completed by the patients was 16.5 seconds. The mean time taken by the university students was 6.9 seconds. The difference gave a "*t*" of 11.3 which was significant well beyond the .001 level of confidence. On the level tests, the mean score for the patients was lower than that of the students but the difference did not reach statistical significance.

The clear cut character of the results is further indicated when the patient's raw scores were transformed into standard scores on the basis of the student data. This enabled Ogilvie to compare the performance of each patient with that of the student controls, on each of the tests. In every case but two, the patients did relatively better on the Level test than they did on the speed test. The generality of this finding was further confirmed by Eysenck in a parallel study [2], who reports results of giving the speed and level tests to non-university normals, neurotics and psychotics. He found that there was no difference between the three groups in the level of problem solving ability. The psychotics, however, were significantly slower than the other two groups. The normals and neurotics did not differ significantly in speed.

Conclusions

First, we have the incidental but useful finding that, compared with nine year old children, adults find it difficult to comprehend simple material which they read aloud. Whatever may be the cause of this difficulty, it argues for the need for considerable caution in the interpretation of an adult's score on a test which has been standardized on children.

Second, we have a finding of general importance. This finding directs attention to a form of impairment which tends to be ignored: a lowering of speed of cognitive function independently of level. Current concepts of deterioration usually imply that mental disorder produces a lowering in the level of cognitive ability. The vocabulary test is used to estimate past level, and non-vocabulary tests are used to estimate present cognitive level. The discrepancy between these two measures is usually taken as an index of the patient's deterioration. This kind of index was first developed by Babcock. It should be remembered, however, that Babcock also suggested that the cognitive deterioration of psychiatric patients might consist of a discrepancy between present level and present speed of function and she expected her index to measure this discrepancy as well. It has been demonstrated elsewhere [5, 10] that Babcock did not provide unequivocal evidence for her theories. The evidence reported in this paper is in effect, the beginning of the provision of such evidence.

Third, it would seem that the outcome of the investigation reported in this paper, confirms the soundness of the methodological conceptions described at the outset. The investigation of a single aspect of psychological function on a single case does not necessarily bog down the investigator amongst a mass of unrelatable observations, so long as the investigation is linked from the very first with the aim of explanation, and so long as the attempt is made to check every explanation by controlled observation. With this precaution, the investigation of the single case can lead to discoveries of general significance.

This conclusion does not imply that the psychologist must always start from scratch. It is obviously better to take validated descriptive generalizations about psychiatric patients as starting points of investigations, if they are available. Even then, it would probably be necessary to return frequently to the single case investigation, at various stages of the development of a research project. The validity of this last proposition, however, still remains to be tested.

Finally, we have to consider how far we have gone in the explanation of the initial observation: the patient's inability to comprehend simple material when read aloud. It would seem that in the first place this difficulty is shared by many normal people of above average ability. It does not, however, exist to the same degree among them as it did in our patient.

The greater degree of difficulty shown by our patient may be due to two reasons. First, the patient's relative slowness of reading might have made it more difficult for him to comprehend what he read. A certain minimum rate of reading might be necessary to facilitate comprehension. In the second place, the patient's rate of comprehension might have been too slow for even his rate of reading aloud. The need to test these two hypotheses is emphasized by the fact that the patient showed no difficulty of comprehension in the silent reading test, providing he was given enough time. This might mean, in accordance with the first theory, that he had in this situation the opportunity of going over the material more than once, or, in accordance with the second theory, the opportunity of reading at his own rate of comprehension. The extent to which the two explanations were correct could only be determined by experiment, and this we did not have the opportunity of doing.

The more general complaint of the patient, his inability to complete the reading of a book, has not of course been finally explained. It might well be that his slowness of performance made it difficult for him to comprehend the main burden of what he was reading and, therefore, he became lost and had to give up the task. A detailed investigation of his reading activity would have been necessary to clear up this matter. It suffices for the purpose of this paper, merely to note that despite the final incompleteness of the investigation, we had arrived at generalities of some importance, both to the patient himself and to clinical psychology in general.

Summary

This paper gives an account of the application of the single case method in the opening up of a field of research in clinical psychology. The investigation results in drawing attention to, and providing confirmation of, one of Babcock's conceptions of deterioration, which has recently suffered some neglect. This is to the effect that cooperative psychotic patients may be very slow in the carrying out of cognitive operations without suffering much loss in the level of difficulty they attain when given sufficient time.

References

1. Burt, C. *Mental and scholastic tests.* London: Staples Press, 1947.
2. Eysenck, H. J. *Differential cognitive tests.* In: U. S. Navy Contract No. N625585-56, 1953.
3. Furneaux, W. D. Some speed, error and difficulty relationships within a problem solving situation. *Nature*, 1952, 170, 37.
4. Furneaux, W. D. *A note on the 'Nufferno' tests of inductive reasoning abilities, including summaries of such norms as have so far been established. Part 2. Speed. The S. E. Tests.* 1953 (Unpublished paper).
5. Nelson, E. H. *An experimental investigation of intellectual speed and power in mental disorders.* Ph.D. Thesis, Univ. London, 1953.
6. Ogilvie, B. C. *A study of intellectual slowness in schizophrenia.* Ph.D. Thesis, Univ. London, 1953.
7. Schonell, F. J. and Schonell, F. E. *Diagnostic and attainment testing.* London: Oliver & Boyd, 1950.
8. Shapiro, M. B. Experimental studies of a perceptual anomaly. I. Initial experiments. *J. Ment. Sci.*, 1951, 97, 90-110.
9. Shapiro, M. B. Experimental studies of a perceptual anomaly. II. Confirmatory and explanatory experiments. *J. Ment. Sci.*, 1952, 98, 605-617.
10. Shapiro, M. B. Use of psychological tests for the differential diagnosis of early dementia. *Proc. Roy. Soc. Med.*, 1952, 45, 375-376.
11. Shapiro, M. B. Experimental studies of a perceptual anomaly. III. The testing of an explanatory theory. *J. Ment. Sci.*, 1953, 99, 394-409.

PREDICTIONS BY MEANS OF TWO PROJECTIVE TESTS OF PERSONALITY EVALUATIONS MADE BY PEERS[1]

HERBERT ZIMMER

Air Force Personnel and Training Research Center
Maxwell Air Force Base, Alabama

Problem

This study sought to assess the validity of two projective instruments, a sentence completion test and human figure drawings, under conditions which approximate their clinical usage as closely as the application of rigorous experimental controls would permit. Specifically, the experimental conditions paralled common clinical practice in two important ways: (a) specific personality characteristics were described and predicted from the tests, and (b) these predictions were made on the basis of an over-all impressional inspection of the test record, rather than after detailed item-by-item scoring and analysis.

Most validation studies have been concerned with the identification of clinical groups, or with the prediction of broad, general characteristics, such as adjustment, improvement, college grades, field dependence, etc. Only a very few validation studies on either of these two tests have attempted to predict specific personality traits [10, 12]. In both of these studies there is a question of possible contamination of the criterion with the predictor.

Validation studies on the sentence completion test and on human figure drawings are far too numerous to be reviewed here. However, many of these studies are covered by one of several available surveys of the literature [1, 2, 3, 7, 8].

Procedure

Criterion scores on eight personality dimensions for 73 subjects were derived from the rankings of each subject by his fellow bomber crew members. These rankings were normalized, pooled, and corrected for differences in error variance due to differences in the size of crews. Three judges attempted to predict these criterion scores by means of a "blind," over-all evaluation of two projective tests judged independently, a sentence completion test and human figure drawings.

Subjects. The subjects of this study were 73 members of nine separate air crews of the U. S. Air Force. The crews ranged in size from five to ten members. The length of time for which each subject had known his fellow crew members ranged from one to 24 months, with the median at 7.25 months.

Peer Ranking. The subjects were given a booklet for ranking their fellow crew members and themselves on eight personality dimensions. The first page of this form contained explicit directions to the subjects. Subsequent pages listed the personality dimensions, one dimension to a page. The top of each page listed one extreme of a dimension, the bottom the other, each followed by three behavioral definitions. The space between contained a vertical column of numbers from 1 to 11. Next to the numbers the subjects wrote the letter abbreviations of the airplane positions of their fellow crew members, in the order in which they fell on the characteristic described on the particular page. The tops of the pages began alternately with the socially desirable and undesirable extremes of the eight dimensions, in order to minimize halo.

Sentence Completion Test. When the subjects had finished the peer ranking, they were given a sentence completion test especially constructed for this experiment. For each personality trait under study, the test contained four sentence completion items, which were intended to elicit material relevant to the specific trait. In addition, there was one general item, making a total of 33 items.

[1]The author wants to thank Dr. Edward E. Cureton for planning the statistical procedures employed in this study, Messrs. William A. Kobzeff, Marion F. Sobczynski and Louis M. Herman for serving as judges, and Harry M. Henkin and Stephen F. Zolczynski for supervising the statistical computations.

Human Figure Drawings. After the subjects had finished the sentence completion test, they did two drawings. The last two pages of the booklet were two blank sheets of bond paper. The subjects were asked to: "Draw the picture of a person on the first of these sheets." When they had finished this drawing they were then told: "Now, draw the figure of a person of the sex opposite to your first drawing. If you have drawn a man, now draw a woman. If you have drawn a woman, now draw a man." All questions were answered in the standard, noncommital manner for administering projective tests.

Judges. In order to insure rankings by the judges comparable to those made by the crew members, the judges (a) ranked each crew as a unit, and (b) ranked the crew members on the same traits, using identical definitions of those traits, as the crew members had done earlier. The judges were instructed to attempt to rank each subject, by means of the test, in the same way the subject's fellow crew members had most likely ranked him. Each judge ranked all 73 subjects on one trait, and by means of one test, before proceeding to the next trait. This was done to maintain uniform standards in evaluating all crews on the same trait and with the same test. The eight personality traits and the two tests were taken in random order by the judges.

When making their rankings on a trait on the basis of the sentence completion test, the judges were asked to read the whole test over first, then evaluate the four test items which were intended to elicit material relevant to the trait, and last to modify their judgment by whatever other relevant information might appear in the whole test record, before ranking the subject. In ranking the human figure drawings, the judges were instructed to be guided first by their over-all impression of the drawings, and if this was inadequate for a decision, to consider various details of the drawings. For the guidance of the judges a list was provided for each of the eight personality traits, which pointed out those aspects of the drawings most relevant to the evaluation of each trait. Whenever the judges were uncertain about the exact rank

TABLE 1. MEDIAN RELIABILITY AND MEDIAN VALIDITY OF SENTENCE COMPLETION TEST AND HUMAN FIGURE DRAWINGS BY TRAITS

Personality dimension	Criterion reliability	Interjudge reliability SCT	Interjudge reliability HFD	Interteat reliability	SCT prediction of criterion	Pooled SCT prediction	HFD prediction of criterion	Pooled HFD prediction	Prediction of criterion by both tests
Boastful-modest	.84	.32	.53	.00	.05	.05	.15	.15	.14
Self-denying self-indulgent	.81	.32	.31	.03	.08	.20	.08	.19	.14
Tense-relaxed	.79	.09	.18	−.04	.10	.10	.15	.18	.21
Assertive-passive	.84	.23	.44	.09	−.03	−.09	−.08	−.09	−.07
Withdrawn-sociable	.82	.45	.39	.08	.05	.07	.06	−.08	−.01
Adaptable-rigid	.79	.03	.26	−.14	.18	.37	.14	.19	.23
Unskilled in conversation-skilled in conversation	.82	.39	−.12	.15	.32	.44	.05	.10	.24
Self-reliant-dependent	.82	.38	.41	−.05	.20	.30	.21	.21	.32
Over-all median	.82	.32	.32	.04	.12	.15	.07	.16	.17

Note: For 71 df an r of .23 deviates significantly from zero in the plus or minus direction at the .05 evel of confidence, .30 at the .01 level, and .38 at the .001 level.

of one or more subjects in a crew, they were encouraged to work out the ranking by applying the method of paired comparison.

Criterion Scores. A criterion score for each subject, on each of the eight personality dimensions, was derived from the pooled rankings of his fellow crew members, with his self-ranking excluded. Following the procedure described by Cureton[4], the rankings were normalized, pooled, and corrected for differences in the size of the error variance, due to differences in the size of crews and consequently in the number of raters. Thus, each subject obtained an average score on each of the eight traits, which represents the combined judgments of the subject's personality by his fellow crew members. Since these scores were normalized and corrected for differences in the size of the crews, they are comparable, across crews, for all 73 subjects. Judge's rankings were similarly normalized to permit the application of Pearson product-moment correlations to all subjects as one group.

Results

The median Pearson correlations given in the tables present the results of this study in summary form. Table 1 gives the median correlations for the eight personality dimensions, and Table 2 shows similar information for the three judges.

Reliability of Criterion Scores. The reliability coefficients of the criterion scores reported in Table 1 were obtained by an extension of the method described by Cureton[4]. The reliability of the mean scores (pooled rankings) of the N subjects, which constitute the criterion scores, is given $r_k = \frac{k r_1}{1 + (k-1) r_1}$ [5]. This boosts the reliability of the single ranking by the average number of rankings per subject, thus indicating the reliability of the pooled rankings which are the criterion scores.

The resulting reliability coefficients suggest that agreement amongst crew members about the personality characteristics of one of their peers was sufficiently great to lend their pooled rankings some measure of consensual validation. This would seem to make the criterion measures adequate to our purpose.

Table 2. Median Reliability and Median Validity of Sentence Completion Test and Human Figure Drawings by Judges

	Judge		
Reliability	A with B	A with C	B with C
Interjudge reliability SCT	.21	.38	.36
Interjudge reliability HFD	.27	.42	.32
	Judge		
Validity	A	B	C
SCT prediction of criterion	.10	.10	.21
HFD prediction of criterion	−.02	.12	.16
Prediction of criterion by both tests	.04	.18	.22

Note: For 71 df an r of .23 deviates significantly from zero in the plus or minus direction at the .05 level of confidence, .30 at the .01 level, and .38 at the .001 level.

Interjudge Reliability. The second and third columns of Table 1 give the interjudge reliability of each test by personality trait. It may be noted that for human figure drawings interjudge reliability for conversational skill is completely lacking. Table 2 shows interjudge reliability by judges. Judge C shows greater agreement on both tests with both other judges, than Judge A and B do with each other. In evaluating

these reliability figures three factors which tend to lower reliability should be considered: (a) the judges were evaluating highly specific traits, and not an all-inclusive factor like "adjustment", (b) the judges did not employ any systematic scoring procedure, but used a "global" type of judgment, and (c) they had no common previous experience.

Intertest Reliability. Each judge made two predictions for each individual on each trait by means of the two separate instruments. The fourth column of Table 1 illustrates the virtual absence of any relationship between the two sets of predictions made on the basis of the two tests.

Prediction of Criterion. The median correlations resulting from the independent predictions of the criterion by Judge A, B, and C, using the sentence completion test, are given under the heading "SCT prediction of criterion," by traits in Table 1, and by judges in Table 2. Under the heading "HFD prediction of criterion" the same information is reported for human figure drawings. The pooled predictions of the three judges, obtained by correlating the algebraic sum of the normalized rankings of one subject by all three judges with the subject's criterion score, are shown in Table 1 under the caption "Pooled SCT prediction" for the sentence completion test, and under "Pooled HFD prediction" for human figure drawings. The median correlations obtained by predicting each subject's criterion score with the algebraic sum of each judge's normalized ranks of the subject's sentence completion test and human figure drawings appear under the heading "Prediction of criterion by both tests" in Tables 1 and 2.

Accuracy of prediction varies not only from judge to judge, but also with the trait predicted, and with the test used. Furthermore, some traits are predicted better by one instrument than by the other. (a) The predictive ability of Judge C is consistently superior to the other two judges, as illustrated in Table 2. (b) For two traits, assertive-passive and withdrawn-sociable, prediction is uniformly poor on both instruments. (c) There is a slight tendency, somewhat obscured by the reporting of median correlations, for the sentence completion test to be a little better predictor than human figure drawings. This, despite the fact that all traits, except conversational skill, were selected as being the kind of personality variable which is frequently derived from human figure drawings. (d) Conversational skill is predicted better by the sentence completion test, as might be expected, while the trait boastful-modest is predicted somewhat better by human figure drawings.

Both combining the evaluation of one test by three judges and combining one judge's evaluation of both tests, result in some improvement of prediction. These combinations tend to compensate for weaknesses amongst either the judges or the instruments, when the elements they are combined with are stronger predictors.

Discussion

Agreement amongst crew members ranking each other is apparently adequate to give the criterion measures acceptable reliability. The pooling of rankings, making each criterion score a mean of scores, and the correction of these scores for differences in the size of error variance, have combined to restrict the range and variability of criterion scores. The artificially narrowed range and variability of the criterion scores tends to lower the correlations of other factors with these scores.

The method used in this experiment, that of "blind" global (no scoring or detailed analysis) evaluation of projective tests, expressed as rankings or ratings on a number of fairly specific personality dimensions was most closely approximated by the Michigan assessment of clinical psychology students [10, 18]. Surprisingly enough, the correlations reported from the use of the sentence completion test in that study are of a similar order of magnitude as those found in the present experiment. Kelly and Fiske [10] report a median correlation of .19 for the sentence completion test against 12 criterion variables. Using 15 different criterion traits from the same study, Samuels [18] gives a median correlation of .28 for the sentence completion test, used

by six independent judges on 2 to 37 subjects per judge. He reports median intertest correlations of sentence completion test ratings with Rorschach ratings of .05, with TAT ratings of .06, and with Bender Gestalt ratings of .08. In a smaller study using the same data on only 20 subjects, four judges, and six traits, Samuels [15] obtained a median interjudge reliability of .35 for the sentence completion test. For the same smaller study, Kelly and Fiske [10] report a median validity for the sentence completion test against four criterion variables of .05.

Hamlin's observation [9] that judges can best deal with bite-size units of intermediate complexity, neither atomistic nor all-encompassing, is probably pertinent here. The apparently poor showing of projective instruments in both the present and the Michigan study may partly be due to the global nature of the evaluations made by the judges. Inasmuch as many clinicians are in the habit of relying on global impressions, especially with the two instruments discussed here, the advantages of somewhat more detailed and specific analysis would appear to merit consideration. This is further emphasized by the success claimed by several studies applying "interpretative" scoring methods to these tests [6, 11, 12, 13, 14, 16].

Summary

This experiment sought to determine the accuracy with which specific personality characteristics can be predicted by means of two projective tests: a sentence completion test and human figure drawings. The criterion scores on eight personality traits for 73 subjects were derived from the rankings of each subject by his fellow bomber crew members. These rankings were normalized, pooled, and corrected for differences in error variance, due to differences in the size of crews. The obtained median reliability of criterion scores was .82. Three judges attempted to predict the criterion scores by means of a "blind," over-all evaluation of each projective test, judged independently. Median interjudge reliability was .32 for each test. The median correlations between judges' predictions and the criteria were .15 for the sentence completion test, .16 for the human figure drawings, and .17 for the judges' pooled evaluation of both tests employed independently. There was a wide range in predictive efficiency with individual r's ranging from −.18 to .44.

References

1. Anderson, H. H., and Anderson, G. L. *Projective techniques.* New York: Prentice-Hall, 1951.
2. Bell, J. E. *Projective techniques.* New York, N. Y.: Longmans, Green, 1948.
3. Brower, D., and Abt., L. E. *Progress in clinical psychology.* New York: Grune and Stratton, 1952.
4. Cureton, E. E. Note on the scaling of ratings or rankings when the numbers per subject are unequal. *Psychometrika,* 1952, *17,* 397-399.
5. Cureton, E. E. Personal communication. August, 1953.
6. Friesen, E. P. The incomplete sentences technique as a measure of employee attitudes. *Personnel psychol.,* 1952, *5,* 329-345.
7. Goodenough, F. L. Studies in the psychology of children's drawings. *Psychol. Bull.,* 1928, *25,* 272-283.
8. Goodenough, F. L. and Harris, D. B. Studies in the psychology of children's drawings: II. 1929-1949. *Psychol. Bull.,* 1950, *47,* 369-433.
9. Hamlin, R. M. The clinician as judge: implications of a series of studies. *J. consult. Psychol.,* 1954, *18,* 233-238.
10. Kelly, E. L., and Fiske, D. W. *The prediction of performance in clinical psychology.* Ann Arbor: Univ. of Michigan Press, 1951.
11. Morton, R. A controlled experiment in psychotherapy based on Rotter's social learning theory of personality. Unpublished doctoral dissertation, Ohio State Univ., 1949.
12. Rohde, A. R. Explorations in personality by the sentence completion method. *J. appl. Psychol.,* 1946, *30,* 169-181.
13. Rotter, J. B., and Willerman, B. The incomplete sentence test as a method of studying personality. *J. consult. Psychol.,* 1947, *11,* 43-49.
14. Rotter, J. B., Rafferty, J., & Schachtitz, E. Validation of the Rotter incomplete sentences blank for college screening. *J. consult. Psychol.,* 1949, *13,* 348-356.
15. Samuels, H. The validity of personality-trait ratings based on projective techniques. *Psychol. Monogr.,* 1952, *66,* No. 5 (Whole No. 337).
16. Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., and Wapner, S. *Personality through perception.* New York: Harper, 1954.

DIRECTIVE THERAPY WITH FEMALE JUVENILE DELINQUENTS

SIDNEY KOSOFSKY

N. J. State Home for Girls
Trenton, N. J.

Problem

The purpose of this paper is to describe the results of the application of directive therapy in a treatment program instituted in a State operated correctional institution for female juvenile delinquents. Girls admitted to this institution, after spending a few weeks in a reception cottage, are assigned to different cottages. Experience had shown that there were always some girls who found it extremely difficult to adjust for a variety of reasons and manifested their maladjustment either by repeated attempts to elope and/or by persistent and intense outbreaks of hostility. To meet this problem, a Treatment Unit was established. The most overtly maladjusted girls were segregated in this Unit in order to (a) relieve the pressure which they exercised on other girls as well as on supervisors; and (b) in order to give these girls more time and individual attention.

Subjects

The Treatment Unit was put into operation in the Fall of 1953. During the course of the first year, 53 girls were admitted to this Unit. Since 20 of the 53 were still in residence at the end of the first year and since this paper covers only the first year of work, the 20 in residence at the end of the first year will not be taken into account. Of the remaining 33, 5 were transferred from this institution to another correctional institution; 4 were transferred to a state hospital; 1 eloped; 1 was recalled by her committing judge; and 22 were transferred from the Treatment Unit to other cottages within the same institution.

Of these 22 girls, 3 were classified as of borderline intelligence, 6 were classified as of dull normal intelligence, 10 were classified as of average intelligence, and 3 were classified as of bright intelligence. The average length of stay for these girls in the Treatment Unit was 123 days with a range of from 50 days to 242 days.

Rationale

In undertaking therapy with this group of delinquents, the writer chose to adopt the directive method of psychotherapy advocated most strongly by Thorne[2-9]. While Thorne indicates that this general approach may be contraindicated with extremely negativistic personalities, the writer, in making this the treatment of choice, was guided by several considerations: (1) in making diagnostic studies of these cases, the writer had come to the conclusion that these girls were never able to impose inner controls on their impulses and behavior and could not be reasonably expected to do so in a relatively short time if the type of therapy and treatment to which they were subjected were to be permissive; (2) on the basis of previous experience with female delinquents, the writer had come to the conclusion that a psychologist—who is regarded as an authority figure despite his attempts to avoid being identified as such—becomes a person of little or no importance in the eyes of these girls if they feel that he has no authority; (3) since the entire institution is run along authoritarian lines and since the whole procedure of commitment to the institution was done from an authoritarian point of view (police, courts, etc.), it was felt that a different approach would be more confusing than helpful to the girls involved.

As conceived by the therapist, a girl's experience in the Treatment Unit was to be part of a re-educative process. The focal point of this process would vary from girl to girl—depending on what the particular chief problem was—but certain general principles were applicable to every girl. Briefly, the factors common to each girl included: (1) the *sine qua non* of establishing a good relationship with her; (2) the affording of opportunities, as often as necessary, for the release of emotional tension;

(3) an attempt to have each girl gain some insight even if this gain were a modest one; (4) an opportunity to work through—either in actual behavior or on a verbal level—whatever insight may have been gained; (5) a directive emphasis on the particular type of behavior which, in the therapist's opinion, was most basic in causing the girl's maladjustment.

Procedure

The attempt to put these principles into operation began with the very first interview. Here, each girl assigned to the Treatment Unit by the staff was given an opportunity to discuss her pattern of behavior, her problems, her difficulties, etc. In a supportive fashion, the girl was brought around by the therapist's questions and discussion to the point where, having verbalized the idea that she had been "bad", the reasons for her assignment to the Treatment Unit were discussed with her. An attempt was made to show her that this was not a punitive measure; that, on the contrary, it could be used by her as an opportunity to effect a change for the better in her attitude and pattern of behavior. She was told that it would be expected that she would behave in an acceptable manner, that each supervisor in the cottage would write a report to the therapist of her behavior each day but, at the same time, she was reassured of the therapist's interest in her and of his sincere desire to be of help to her if she were willing to accept such help.

Because, after each visit to the Treatment Unit, the writer made brief notes as to whom he saw, what transpired, etc., he is able to note the following as the reactions of the girls to the first interview:

Table 1. Reactions to First Interview[1]

Personality Reactions	Cases	Percent
1. Very cooperative. Reaction positive initially and remains positive throughout	2	9
2. Cooperative with reservations. Tentatively accepts therapist. Affective reaction friendly.	3	14
3. Polite on surface. Neutral and somewhat indifferent.	4	18
4. Openly resistive at first. Sometimes cries in anger. Finally accepting and amenable.	2	9
5. Poorly concealed negativity. Openly suspicious of therapist. Reaction predominantly negative.	5	23
6. Openly negativistic and resistive. Rejects efforts to establish rapport. Says she is "bad" and will remain that way despite any attempts to have her change.	5	23
7. Initially accepting and amenable; then becomes resistive and negativistic.	1	4
	22	100

[1]Table 1 uses the same type of classification as was used by Thorne (7, p. 76). Some modifications of the personality reactions described by Thorne are necessary because none of the girls in this particular situation was in a position to seek treatment on her own initiative. The personality reactions are, therefore, described somewhat differently from, but along the same general lines as, those of Thorne.

Each girl was seen by the therapist on the average of three times a week either in a formalized individual interview or when she was part of a group of girls in a group session. Any girl who was reported for misbehavior was always seen by the therapist immediately upon receipt of the report. The particular incident was gone over in great detail and an attempt was made to use it as a teaching device. If disciplinary action was indicated, the therapist made it very clear that, though he did not administer the discipline itself, he was responsible for its administration. (A disciplinary measure always took the form of deprivation of one or more privileges). An attempt was always made—usually successfully—to have the girl arrive at a point where she could accept the disciplinary measure since it was felt that only by her acceptance of it could it serve its function maximally. When, as sometimes hap-

pened, the disciplinary measure evoked hostility on the part of the girl, the therapist accepted the hostility and tried either to blunt it or to redirect the engendered emotion along the lines of increasing motivation and/or mobilizing one's resources for another attempt to exercise self-regulation. Every effort was made to treat each girl in a consistently fair and honest manner. The principles of suggestion, persuasion and advice as formulated by Thorne [7, p. 73f.] were always borne in mind. Attempts at influencing each girl were kept at a minimum, were made in impersonal terms, and the element of her own responsibility was constantly stressed. It was part of the deliberate therapy plan to avoid making such attempts too early, too often, or in a dictatorial fashion.

A review of the therapist's notes gives a fairly clear, though somewhat incomplete picture of the topics which were discussed and of their relative frequency. These are shown in Table 2.

TABLE 2. TOPICS DISCUSSED IN INDIVIDUAL AND GROUP THERAPY SESSIONS

Topics	Individ. Sessions	Group Sessions
1. Importance of getting along with other people; necessity to trust others; importance of individual friendship with others; fighting.	28	18
2. The role of attitude in determining behavior; perseverance in motivation to improve one's self.	19	5
3. Necessity of accepting authority; controlling one's hostile impulses; dealing with the desire to elope from the institution.	21	2
4. Parent-child relationships in all its ramifications.	18	3
5. Reassurances; usually given directly by the therapist and dealing with such matters as "being crazy", sex questions, dreams, etc.	18	3
6. Planning for parole; facing one's problems realistically.	10	2
7. Lying as unacceptable behavior.	7	1
8. Controlling one's temper.	7	6
9. Stealing as unacceptable behavior.	6	2
10. Stubbornness as a personality trait.	3	1
11. Sex- dealing with both homosexuality and heterosexuality.	5	4
12. Assuming responsibility for one's welfare and well being—including such matters as mutilating one's self, suicide, destroying property, etc.	4	2
13. Developing ability to think critically; overcoming high suggestibility.	4	1
14. Race relations.	2	4

It should be noted that the single most powerful motivating factor in the girls' behavior was their desire to "get open". In using this phrase, the girls were expressing both their immediate and secondary goals, i.e., to be transferred from the Treatment Unit to another cottage and to be released from the institution on parole. It was this factor which was all important in determining what topics they discussed with the therapist, what attitudes they adopted, and what goals they articulated. However, it may be noted that the therapist used some directiveness—in a subtle fashion—in bringing to the fore certain topics which sometimes did not seem to be directly relevant to the girls' central purpose. In introducing such topics, the therapist was very much influenced by the work of the Gluecks [1] who have emphasized a cluster of five personality traits which are very frequently found in delinquents: adventurousness, extraversion in action, stubbornness, suggestibility, and emotional instability. It was because the therapist's experience with female delinquents tended to confirm the findings of the Gluecks with male delinquents that he introduced discussions on such topics as stubbornness, overcoming high suggestibility, etc. Very frequently, the introduction of these topics and their elaboration and discussion resulted in the gaining of some new insight on the part of one or more of the girls.

As part of the treatment program, some work was done with teachers and cottage supervisors who came into contact with the girls. In private discussions and in

more formal meetings of these groups, the therapist undertook to explain the directive technique—with some of its indications and contraindications—and his rationale for using it.

Termination

A girl's release from the Treatment Unit was wholly dependent on her attitude towards others and on her overt behavior. Following a period of stabilization, during which a girl behaved in an acceptable fashion because of what seemed to be a change for the better in her attitude, she was transferred to another cottage. It was recognized that the combination of controls and support which had been given to her in an intensive fashion in the Treatment Unit could not be abruptly withdrawn without risking a regression on her part to her previous pattern of behavior. For this reason, the therapist made it a practice to see each girl who was released from the Treatment Unit on an average of once a week—sometimes more often if she requested it—in order to give her some directive guidance and counsel. With the passage of time and with the proven ability of the girl to continue to make a good adjustment, she was gradually weaned from the therapist's guidance and counselling. Where the transition from the Treatment Unit to another cottage was too great a step as shown by her regression to her previous unacceptable pattern of behavior, a girl was returned to the Treatment Unit for a second time. Three of the 22 girls fell into this category. During the follow-up period, both group and individual sessions were used and much the same procedure was followed as was the case when the girls were in residence in the Treatment Unit.

Results

Because this article is being written only four months after the end of the first year of the Treatment Unit operation, it is impossible to give any long-term results. Generally, it may be said that the operation of the Treatment Unit was not always a smooth and easy process. Several times girls eloped from the Unit in groups of two or three and, on one occasion, in a larger number. In addition, there were a number of girls who periodically expressed their hostility by destroying furniture, tearing clothes, flooding a room, etc. As might be expected, too, there were some stealing episodes as well as some fights among the girls.

Table 3 gives the results of the first year's operation in summary form:

Table 3. Disposition of Cases Following Directive Psychotherapy.

Result	N
Success in institutional adjustment	19
Failure in institutional adjustment	3
Success on parole	13
Failure on parole	3
Not yet paroled at the time of writing	6

That there is no perfect correlation between a successful institutional adjustment and success on parole is indicated by the fact that two of the three girls who failed on parole had made a successful institutional adjustment while two of the girls who had failed to make a good institutional adjustment were succeeding on parole at the time of the present writing.

Since the girls who were in residence in the Treatment Unit were those who had been the most maladjusted behaviorally, there was no comparable group which could serve as a control. However, it was the consensus of the staff that, both for the advantages which it offered to this type of girl as well as for the advantages which accrued to the institution as a whole, the program was very worthwhile and, at the time of the present writing, it is being continued.

Summary

One year's experience in segregating the most serious behavior disorders in a population of female juvenile delinquents and the use of directive psychotherapy with both individuals and groups is described. Some of the characteristics of the group, the rationale of the procedure, the technique used, and the results are given.

References

1. Glueck, S. and Glueck, E. *Unraveling Juvenile Delinquency.* New York: Commonwealth Fund, 1950.
2. Thorne, F. C. Directive Psychotherapy: III. The psychology of simple maladjustment. *J. clin. Psychol.*, 1945, 1, 228-240.
3. Thorne, F. C. Directive Psychotherapy: VII. Imparting psychological information. *J. clin. Psychol.*, 1946, 2, 179-190.
4. Thorne, F. C. Directive Psychotherapy: VIII. Personality integration and self-regulation. *J. clin. Psychol.*, 1946, 2, 371-383.
5. Thorne, F. C. Directive Psychotherapy: IX. Therapeutic use of conflict. *J. clin. Psychol.*, 1947, 3, 168-179.
6. Thorne, F. C. Directive Psychotherapy: XII. The client's Weltanschauung. *J. clin. Psychol.*, 1947, 3, 277-286.
7. Thorne, F. C. Directive Psychotherapy: XIV. Suggestion, persuasion and advice. *J. clin. Psychol.*, 1948, 4, 70-82.
8. Thorne, F. C. Directive Psychotherapy: XV. Pressure and coercion. *J. clin. Psychol.*, 1948, 4, 178-188.
9. Thorne, F. C. Principles of directive counselling and psychotherapy. *Amer. Psychologist*, 1948, 3, 160-165.
10. Thorne, F. C. *Principles of Personality Counseling.* Brandon, Vermont: Journal of Clinical Psychology, 1950.

THE DISCRIMINANT VALUE OF MINNESOTA MULTIPHASIC PERSONALITY INVENTORY ITEMS IN PERSONALITY DISORDERS*

REUBEN S. HORLICK

Audiology and Speech Correction Center
Walter Reed Army Medical Center
Washington, D. C.

PROBLEM

A frequent problem at psychiatric clinics and hospitals has been the need to differentiate the neurotic from the psychotic, and the neurotic from the normal. Meehl[1] has indicated that a chief difficulty in analyzing the Minnesota Multiphasic Personality Inventory is distinguishing between neurotic and psychotic profiles. Rubin[2] believes, after suggesting that the failure of the MMPI scores to show a significant agreement with psychiatric diagnoses may be due to the lack of reliability of such diagnoses, that analysis of the individual items will result in information which will aid in the understanding of the dynamics underlying the maladjustment in behavior.

Studies on individual items have been made by several investigators. Manson[3] conducted a study on a group of alcoholics and Brozek and Erickson[4] investigated the effect on an experimental group on a semi-starvation diet. Winne[5], in an attempt to develop a Scale of Neuroticism, analyzed the 117 items composing the Hypochondriasis, Depression and Hysteria Scales of the MMPI to select "items which would distinguish between normals and neurotics." Wauck[6] investigated the Schizophrenia Scale to determine whether this scale by itself is a valid indicator of psychosis. In both studies, however, comparisons with other psychiatric groups were not made.

The purpose of this study is to investigate the differential value of MMPI items in discriminating between two groups of psychiatric patients; more specifically, to determine whether certain of the items will separate or distinguish a group of anxiety-neurotic patients from a group of paranoid-schizophrenic patients and, as a corollary problem, to determine whether any items should be added or subtracted from the scoring keys already in use.

METHOD

The subjects in this study were male, neuropsychiatric Army patients including a group of 25 neurotics, a second group of 50 schizophrenic patients and a third group of 25 out-patients from the Neuro-psychiatric Out-patient Clinic.

The mean ages were found to be 27.06 (SD 5.75) for the paranoids, 28.28 (SD 4.91) for the neurotics, and 26.12 (SD 5.76) for the out-patient group. The mean number of years of schooling and the standard deviations for each of the three groups were computed. The paranoid patients, on the average, completed 11.08 years of schooling with a standard deviation of 2.53; the neurotics showed 10.88 years (SD 2.72); and the out-patients' education average was 10.92 years (SD 2.36). The critical ratios of the mean differences between any two groups were determined and found to be below the five per cent level of confidence.

PROCEDURE

The first step in the study was a tally count of each item in the MMPI answered critically according to the MMPI scoring keys. Answers to questions for which scoring keys did not exist unfortunately were not recorded and therefore were not available for item analysis. Based on these numbers, the percentages of each group answering each of the scored items critically were obtained. For example, for item

*We wish to express our appreciation to Hewitt Cochran who did the laborious statistical analysis of the data.

A-1—"During the past few years I have been well most of the time"—22 percent of the schizophrenics, 60 percent of the neurotics, and 20 percent of the out-patient group answered this item critically, i.e., false.

The critical ratios for each item were obtained by using Edgerton's Chart [7] and the formula for determining the standard error of the difference,

$$SE_{Diff} = \sqrt{SE^2_{m_1} - SE^2_{m_2}} \text{ and } CR = \frac{D}{SE_{Diff.}}$$

The "*t*" test method was used to examine each critical ratio for level significance.

Results

Table 1 presents a breakdown of the number of items which differentiate the sample of paranoid-schizophrenics from the anxiety-neurotic and the out-patient groups. Thus, 42 scored items tended to differentiate the schizophrenic patients from the other two groups. Six of the items are "low" percentage items, i.e., scored least frequently when compared with the other two groups. These are A-45, C-18, D-10, F-33, G-21, and I-31.

Table 1. Number of MMPI Items Differentiating Each Group of Patients

Group	No of Items	High*	Low*	Level of Significance			Scale on Which Items Occur											
				.01	.02	.05	L	K	F	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma
Paranoid	42	36	6	32	3	7	3	6	8	0	3	6	10	10	6	1	11	4
Neurotic	69	44	25	43	7	19	2	4	2	13	11	22	11	13	4	17	14	11
Out-Patient	53	22	31	26	9	18	3	4	5	3	6	6	7	9	9	9	13	10

*High indicates the number of items on which the given sample of patients rated highest percentagewise; Low indicates the number of items on which the sample rated the lowest percentagewise.

Theoretically, it would appear that "low" items, if used, should follow the McKinley and Hathaway Manual [8] and be scored as blank items, designated on the scoring keys as "O" instead of as "X". Only one item, C-18, is so designated according to the manual. Another item, J-41, while assigned an "O" score in the manual, is not answered at the five per cent level of significance by any of the three groups sampled. It may be a valid "lie" item but its use in the Paranoia, Psychasthenia and Schizophrenia Scales may not be warranted if it is not a discrimination item.

While this study confirms that item C-21 differentiates the schizophrenics from the others at the one per cent level, it would be more appropriate to attach a blank or "O" designation to the item rather than assigning the "X" score as now provided by the MMPI Schizophrenia Scale scoring key. The remaining four items, A-45, D-10, F-33, and I-31, which should belong in the Schizophrenia Scale with an "O" designation, do not appear on the present scoring key. Such inclusion, however, would have to wait on a much more comprehensive study with other groups of patients, as well as normal subjects.

Of the 36 "high" percentage items, i.e., occurring most frequently percentagewise when compared with the other two groups, only six (A-9, E-12, G-34, G-54, H-3, H-16) are at present on the MMPI Paranoia Scale and nine (A-30, C-5, C-18, C-53, D-32, E-17, G-21, G-51, H-12) are on the Schizophrenia Scale. While several of these items are "high" percentage items, the total per cent of the paranoid patients scoring the items was small. For example only 18 per cent scored item C-34 and 22 per cent scored item H-16.

Nevertheless, several of the remaining 36 items (A-5, B-37, B-53) should be reexamined for possible inclusion in the Paranoia or Schizophrenia Scales. Thirty-four

of the 40 items now included in the present Paranoia Scale were not answered critically by the paranoid group and should, therefore, be re-examined to determine whether these items actually are of diagnostic help.

Turning to the study of the neurotic patients, Table 1 reveals 69 MMPI items which differentiate from the paranoid and out-patient groups. An item analysis of these data reveals that 12 "high" percentage items, i.e., items identifying the neurotic group by the frequency of their critical scoring of these items, are at present included in the Schizophrenia Scale. They are (A-26, A-27, A-36, A-37, A-42, A-43, F-35, F-39, F-49, G-20, I-23, I-35). It appears that these items belong more properly in a neurotic grouping rather than in a "psychotic setting." It is not surprising, therefore, to find MMPI profiles occasionally ambiguous in that neurotic profiles show elevated Schizophrenia Scale scores with resultant difficulty in the differential diagnosis. Further study appears warranted, also, with regard to four Paranoia Scale items (A-36, B-54, F-35, and G-12) which were answered critically by a "high" percentage of neurotics.

Of the 44 "high" percentage items, 26 are contained in one of the three scales comprising the neurotic triad; two "low" percentage items (B-40 and E-53) appear on the Hysteria Scale and one "low" item (C-31) is found on the Depression Scale. A mixed number of high and low percentage items appear on the Psychopathic Deviate Scale.

Table 1 shows the distribution of items for the out-patient group. It will be noted immediately that a greater number of low items are represented whereas the reverse was true for the other two groups. This is consistent with the fewer number of items that are marked by non-clinical groups. Of the 22 "high" items, (See Table 2) no special piling up occurs on any one scale, although it may be of significance that the fewest number of items appear on the Hypochondriasis and Schizophrenia Scales. A further analysis shows that 11 items are scored significantly by 50 per cent or more of the out-patients, and of these 11, several appear directly related to the Army environment. For example, 76 per cent of the group indicated that they had "to take orders from someone who did not know as much as (they) did." "High" percentages for the same item were indicated by the other two groups as well (54 per cent and 56 per cent for the schizophrenics and neurotics, respectively). Another item related to the above is G-50, "Some people are so bossy I feel like doing the opposite." Fifty-six per cent of the out-patients agreed with this statement.

TABLE 2. DISTRIBUTION OF "HIGH" ITEMS DIFFERENTIATING OUT-PATIENT GROUP.

Number of Items	Scales on Which Items Occur											
	L	K	F	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma
22	0	2	1	0	3	2	4	4	5	5	2	6

On the other hand, several items which are scored frequently by the out-patients appear on the Psychopathic Deviate Scale. Thus, items E-46, F-8, and I-13 are each scored by at least 60 per cent of the group. The data suggest that the constellation of the group, in an Army setting, is of paramount importance in evaluating their MMPI profiles. Apparently, characteristics of "civilian" populations cannot be utilized as norms for Army populations.

It will be recalled from Table 1 that 13 items from the Schizophrenia Scale significantly differentiated the out-patient group from the other two. Of these, 11 were "low" percentage and quite properly do not apply to this group of patients. However, the other two Schizophrenia Scale items, C-10 and E-38, were high and suggest that these two items are misplaced and do not properly belong in the Schizophrenia Scale. Likewise, four of the Paranoia Scale items were "low" percentage (A-22, C-53, G-55, H-29) and hence, would not as indicated, apply to the out-patients; but five Paranoia Scale items are "high percentage" (D-35, D-50, D-54, F-10, G-50) and are scored by

more than 50 percent of the group. These five items, it would seem, should be re-examined for possible deletion from the scale.

Discussion

While our sampling of patients is too small upon which to base reliable judgments regarding the efficacy of MMPI items to distinguish effectively among various clinical groupings, the evidence produced by this investigation justifies further study with larger numbers of patients in more sharply delineated entities. Because of the differences in the socio-economic and cultural backgrounds as well as in the severity and chronicity of the illness, it is impossible to compare Army populations with the normative or clinical groups used in the original investigations by Hathaway and McKinley. It is clear, however, that the Army environment is in itself a contributing factor in the possible spurious elevation of scores and the resulting distortion or misinterpretation of the data.

Summary and Conclusions

This study investigates the discriminatory value of the MMPI items in personality disorders. Two groups of closed-ward patients, 50 paranoids and 25 anxiety-neurotics, and a group of 25 out-patients from the Neuropsychiatric Service of the Walter Reed Army Hospital were selected for study. All subjects were white, male military personnel and were equated for age and education. Each was administered the Minnesota Multiphasic Personality Inventory and a tabulation was made of the response to each item appearing on any one of the scales. The standard errors of the percentages were computed and the critical ratios of the differences between any two groups were tested for significance. It was found that a number of the items, especially those appearing on the Schizophrenia and Paranoia Scales, are misplaced and that others are non-discriminatory. By re-examining and re-evaluating a number of critical items, it should be possible to revise and modify the MMPI scoring keys, making feasible more representative profiles and more effective differential diagnoses.

Bibliography

1. Meehl, P. E. Profile analysis of the MMPI in differential diagnosis. *J. appl. Psychol.*, 1946, 30, 517-524.
2. Rubin, H. The MMPI as a diagnostic aid. *J. consult. Psychol.*, 1948, 12, 251-254.
3. Manson, M. P. A psychometric differentiation of alcoholics from non-alcoholics. *Quart. J. Studies Alcohol.*, 1948-1949, 9, 175-206.
4. Brozek, J. and Erickson, N. K. Item analysis of MMPI scales in experimental semistarvation. *J. consult Psychol.*, 1948, 12, 403-411.
5. Winne, John F. A scale of neuroticism: An adaptation of the MMPI. *J. clin. Psychol.*, 1951, 7, 117-122.
6. Wauck, L. A. Schizophrenia and the MMPI. *J. clin. Psychol.*, 1950, 3, 279-282.
7. Edgerton, H. A. and Paterson, D. C. Table of standard errors and probable errors of percentages for varying numbers of cases. *J. appl. Psychol.*, 1926, 10, 378-391.
8. Hathaway, S. R. and McKinley, J. C. *The Minnesota Multiphasic Personality Inventory.* New York: The Psychological Corporation, 1946. Rev. Ed.

STUDIES ON THE PHYSIOLOGY OF AWARENESS: THE DIFFERENTIAL INFLUENCE OF COLOR ON CAPILLARY BLOOD-OXYGEN SATURATION

JOHN W. LOVETT DOUST and ROBERT A. SCHNEIDER

Department of Psychiatry, University of Toronto, Canada

Department of Medicine, University of Oklahoma School of Medicine

Problem

The present investigation attempts to assess the physiological influence of color on emotions and awareness. The quality of affect was estimated by a ranking preferance technique and the levels of awareness by their equivalence with blood-oxygen saturation values[8, 9, 10, 11, 12, 13, 14], both variables being considered in terms of response to a sequence of separately presented hues. The effects of color on healthy subjects were compared and contrasted with those on a group of psychiatric patients.

Experimental Methods

The responses of 87 Ss. were investigated. Of this total, 49 were healthy controls (39 males and 10 females) with ages ranging from 18 to 45 years. The remaining 38 Ss. were a group of 21 male and 17 female psychiatric patients aged 16 to 62 years. They included 5 patients with neurosis, 4 with depression, 11 with schizophrenia, 6 with epilepsy and 12 with high-grade mental deficiency. An identical procedure was carried out on each subject in both groups.

The Color Stimuli. These consisted of a series of brilliant poster water colors all of approximately equal intensity and of different hues. Each color filled a circular area 6 cm. in diameter on a white card measuring 20 x 13 cm. The hues were white, a deep royal blue, orange, a medium leaf green, purple, scarlet, black and canary yellow. Thus each complete stimulus consisted of a white or colored circular figure superimposed on a larger background of white. The cards were presented individually in the order listed to the Ss, the white card being repeated at the end of the series. Each card in rotation was turned up from a pile concealed from the S and exposed on the table in front of him. Exposure time was a period of 5 secs. following which an oximetric reading was taken and the card withdrawn. After the individual presentation of each of the stimuli in sequence (and the repeat exposure of the white card), the 7 cards were all spread out together in front of the S and he was requested to select that card which appealed to him most. This card was then withdrawn and the request to rank the remainder repeated, and so on until no further cards remained.

Measurement of Capillary Oxygen Saturation. This was made by the rapid method of discontinuous spectroscopic oximetry. Precise details of the technique are given elsewhere[9, 16]. Essentially it depends upon the fact that the oxyhaemoglobin spectral absorption bands of the capillary blood in the skin vessels disappear following occlusion of the peripheral circulation in a time period which bears a direct relationship to the capillary blood-oxygen saturation levels. The validity of this method has been estimated as $r = .87$ against the concurrent measurement of alveolar oxygen tension[16] and as $r = .92$ against blood-oxygen saturation levels as estimated by a photoelectric[15] oximetric method[9].

Results

In the statements which follow it will be seen that the oximetric measurements are expressed for convenience in terms of "R. T. secs." These represent the actual raw data. Conversion to blood-oxygen saturation may readily be made from them by the use of a regression equation[9].

The Oximetric Response to Color. Table 1 lists the oximetric raw scores for the two clinical groups both before and during exposure to the color stimuli. It shows that the mean resting levels for the control group are significantly higher than those of

TABLE 1. MEAN REDUCTION TIMES AND SIGNIFICANCE OF DIFFERENCES BETWEEN CONTROL AND PATIENT GROUPS

Color Stimulus Presented	Mean Reduction Time (Seconds)			Differences Between Control and Patient Group Means		Difference Between Means of Resting Level and Color Responses.			
						Control Group		Patient Group	
	Total Group (N=79)	Control Group (N=48)	Patient Group (N=31)	t	p	t	p	t	p
Resting Levels	39.709	43.104	34.452	5.024	<.01	—	—	—	—
White I	37.532	40.500	32.935	3.805	<.01	1.704	<.01	.259	>.10
Blue	36.316	39.396	31.548	4.392	<.01	3.779	<.001	1.064	>.10
Orange	35.063	38.521	29.710	4.466	<.01	6.38	<.001	7.33	<.0001
Green	35.759	38.167	32.032	4.228	<.01	9.08	<.001	1.40	>.10
Purple	32.405	35.104	28.226	3.742	<.01	10.30	<.001	3.39	<.01
Red	31.380	33.979	27.355	3.484	<.01	10.863	<.001	3.783	<.01
Black	29.544	31.354	26.742	2.763	<.01	18.68	<.001	4.06	<.001
Yellow	33.177	35.896	28.968	3.555	<.01	6.476	<.001	3.073	<.01
White II	38.494	41.542	33.774	4.807	<.01	1.704	<.01	.259	>.10

the patients. Presentation of the color stimuli was in the order listed in the Table, the white card being repeated at the end of the chromatic series. No significant differences were found ($t = .899$; p is N. S.) between mean resting levels and the presentation of the white card in either control or psychiatric group whether this card appeared early or late in the sequence. Significant differences were found however between the resting oximetric levels and those resulting when all other colors than white were used as stimuli, relative degrees of anoxaemia being associated with each color.

Differential Responses in Controls and Patients. The influence of each hue on the mean resting scores may be calculated from Table 1 which also reveals a diminished reactivity for the psychiatric group as compared with the controls. In order to test the statistical significance of these differences between controls and patients in terms of the difference of the oximetric response to color and the unstimulated resting level, analyses of variance were made. Comparison between differences from resting levels in terms of all the colors revealed that the mean of the total differences for the controls (N = 49) was 7.45 R. T. secs. and for the patients (N = 38) 5.87. The F ratio ($df = 1$ and 782) was 124 (p is .001) but among the groups, differences due to individual hue were not significant (F ratio = .97; $df = 8$ and 782), and there was no significant interaction between clinical categories and colors (F ratio = .55; $df = 8$ and 782).

In terms of relative anoxaemia, even blue, the least stressful stimulus, was actually significant as compared with the resting level for the control group ($t = 3.799$; $p < .001$) and probably so also for the total combined group ($t = 2.425$; $p < .05$), although this difference was not significant when the patients were considered alone ($t = 1.064$; $p > .1$). A similar situation obtained for green. For comparison between the mean resting level and the mean responses to all other colors, the differences were significant for the psychiatric group as well as for the control and total groups.

The Relationship of Differential Anoxaemia to Spectral Position. Inspection of Table 1 also reveals that the mean responses for the chromatic hues may be placed oximetrically along a continuum similar to that seen in the solar spectrum. Excluding black and white whose influences are anomalous, and remembering that our "purple" hue contained very much more saturation with red than does the spectral violet, it may be seen that, for the total combined group, anoxaemia becomes increasingly evident as one proceeds in a sequence of stimuli in the order of blue, green, orange, yellow, purple and red.

Comparison of the differential influence of the three primary hues (of blue, yellow, and red) is made in Table 2. Significant differences were seen for blue *vs.* yellow in the control and total groups, for blue *vs.* red in the same groups, but for yellow *vs.* red the differences were not significant in any of the groups. It may therefore be con-

TABLE 2. COMPARISON OF MEAN DIFFERENCES IN REDUCTION TIME WITH DIFFERENT HUES

Primary Hues Compared	Clinical Groups Studied	Mean Difference (RT SECS)	t	p
Blue *vs.* yellow	Controls	3.500	2.85	$< .05$
	Patients	2.580	.30	$> .05$
	Total	3.139	2.226	$< .05$
Yellow *vs.* Red	Controls	1.917	1.36	$> .10$
	Patients	1.613	.19	$> .10$
	Total	1.797	1.26	$> .10$
Blue *vs.* Red	Controls	5.417	4.18	$< .05$
	Patients	4.193	1.61	$> .05$
	Total	4.936	3.55	$< .05$

cluded that, both for healthy subjects and for many psychiatric patients, the mean oximetric values for the presentation of blue are larger than those for red, while those for yellow tend to be intermediate.

Combining the means of those three colors yielding the highest capillary blood-oxygen saturation values, and comparing these with the means of the three hues giving the lowest values, probably significant (.05) differences were found for all save 4 subjects in both the control and psychiatric groups.

Rank Order of Preference and Its Oximetric Correlates. The rank order of color preference for the total combined group of subjects was found to be: blue, red, green, orange, yellow, purple, black and white. A correlation coefficient was run between the rank orders of preference and the results of spectroscopic oximetry, the differences between the oximetric response to each color stimulus and the resting level being considered. Excluding black and white, *rho* was .43.

Discussion

The color stimuli employed in these experiments were by no means "pure". In an investigation of the affective value of color as a function of hue, tint and chroma, Guilford [6] concluded that hue was the most important factor. All the colors employed in our work were highly saturated, the results of their ranking both by the healthy controls and the patients yielding a rank order entirely comparable with the massed rankings of the 21,060 subjects analysed by Eysenck [2], at least for the blue, red, green, orange and yellow in our series which were also represented in his. This correspondence with a "universal color order" is valuable evidence of the saturation value and distinctive hue of the color stimuli employed here. It is also relevant to note that the stimulus materials did not appear to influence the results of Eysenck's data, all varieties of material being used from colored paper to dyed wools or velvets, lights, crayons or even the names of colors. Similar results were obtained more recently by Granger [4].

Consideration of Table 1 also suggests confirmatory evidence for the conclusions of Kretschmer [7] and Robertson [17] that psychiatric patients display a lessened reactivity in relation to color than do healthy people. Robertson [17] noted a "weaker feeling" among patients as expressed in their own statements, in their readiness to prefer combinations of color, and in the variability of their feelings concerning such color combinations. Our findings clearly demonstrate a diminished oximetric lability among the psychiatric group as compared with the controls and further refute the opinion of Goldstein [3] that colors influence psychotics more than normals.

Summary

(1) The responses of 49 healthy subjects and 38 psychiatric patients suffering from neurosis, depression, schizophrenia, epilepsy and mental deficiency to a series of 7 different color stimuli were assessed in terms of changes in capillary blood-oxygen saturation levels, as determined by spectroscopic oximetry.

(2) For the total group it was found that exposure to any color save white significantly lowered the oximetric score, the differential effect being least marked with hues at the blue-green end of the spectrum and most apparent with those at the red end.

(3) The mean resting oximetric levels of the psychiatric patients were lower than those of the controls, and their reactive lability to different color stimuli was also significantly less than that of the healthy subjects.

(4) The mean rank order of color preference for the combined groups of the present series essentially agreed with that found by other workers. The rank order co-efficient (*rho*) with oximetric scores was .43 for the chromatic hues.

References

1. Birren, F. *Color Psychology and Color Therapy.* New York and London: McGraw Hill, 1950.
2. Eysenck, H. J. A critical and experimental study of color preferences. *Amer. J. Psychol.*, 1941, *54*, 385.
3. Goldstein, K. Some experimental observations concerning the influence of colors on the function of the organism. *Occup. Therap. Rehabilit.*, 1942, *21*, 147.
4. Granger, G. W. Objectivity of color preferences. *Nature*, 1952, *170*, 778.
5. Guilford, J. P. Affective value of color as function of hue, tint and chroma. *Psychol. Bull.*, 1933, *30*, 679.
6. Katz, D. *The World of Color.* London: Kegan Paul, Trench, Trubner & Co., 1935.
7. Kretschmer, E. *Korperbau und Charakter* (20th edn.) Berlin: Springer, 1950.
8. Lovett Doust, J. W. Studies in the Physiology of Awareness: oximetric evidence of the role of anoxia in certain psychiatric states. *Proc. Roy. Soc. Med.*, 1951, *44*, 347.
9. Lovett Doust, J. W. Spectroscopic and photo-electric oximetry in schizophrenia and other psychiatric states. *J. ment. Sci.*, 1952, *98*, 143.
10. Lovett Doust, J. W. Studies on the Physiology of Awareness: oximetric analysis of emotion and the differential planes of consciousness seen in hypnosis. *J. clin. exper. Psychopathol.*, 1953, *14*, 113.
11. Lovett Doust, J. W. and Leigh, D. Studies on the Physiology of Awareness: the inter-relationship of emotions, life situations and anoxaemia in patients with bronchial asthma. *Psychosomat. med.* 1953, *15*, 292.
12. Lovett Doust, J. W. and Schneider, R. A. Studies on the Physiology of Awareness: the effect of rhythmic sensory bombardment on emotions, blood oxygen saturation and the levels of consciousness. *J. ment. Sci.*, 1952, *98*, 640.
13. Lovett Doust, J. W. and Schneider, R. A. Studies on the Physiology of Awareness: anoxia and the levels of sleep. *Brit. med. J.*, 1952, *i*, 449.
14. Lovett Doust, J. W., Schneider, R. A., Talland, G. A., Walsh, M. and Barker, G. Studies on the Physiology of Awareness: the correlation between intelligence and anoxaemia in senile dementia. *J. nerv. ment. Dis.*, 1953, *117*, 383.
15. Millikan, G. A. Oximeter, an instrument for measuring continuously the oxygen saturation of arterial blood in man. *Rev. Scient. Instr.*, 1942, *13*, 434.
16. Ray, G. B. The measurement of the reduction time of blood in the capillaries of the skin. *Amer. J. Physiol.*, 1946, *147*, 622.
17. Robertson, J. P. S. The use of color in the paintings of psychotics. *J. ment. Sci.*, 1952, *98*, 174.

RELATION BETWEEN HISTAMINE TOLERANCE, VISUAL AUTOKINESIS, RORSCHACH HUMAN MOVEMENT, AND FIGURE DRAWING*

STANLEY R. GRAHAM

Creedmoor (N. Y.) Institute for Psychobiologic Studies

Introduction

Previous studies[2, 3] suggest that autokinesis is always accompanied by body tensions in the direction of perceived movement.[1] In the phenomenon of human movement responses elicited during the Rorschach Test there appears to be a similar accompanying physical component, the perception being accompanied by a real or subjectively experienced change in muscular tension[8]. Ach[7], speaking of the factor of awareness in experimental situations, describes tension experiences as a customary

*The author wishes to express his appreciation for the advice and assistance extended by Prof. John G. Rockwell, Chairman of the Department of Educational Psychology of New York University, and Dr. John R. Whittier, Director of Creedmoor Institute, in the preparation of this paper. Dr. Karl Steinman, who constructed the manual for estimating histamine tolerance from human figure drawings, and taught the writer the use of the manual, merits acknowledgment both for his academic contribution and for the substantial amounts of time and energy freely given. The author acknowledges with pleasure the assistance of the Nursing Discipline of the Creedmoor Institute: Mrs. Eileen Jacobi, R. N., Mrs. Eleanor Alesi, R. N., Mrs. Doris Williams, R. N., and Mrs. Evelyn Davis, R. N.

[1]For the purposes of this study the term *visual autokinetic movement* is defined as, "the illusion which refers to apparent movement of a stationary light in the dark."

phenomenon in the upper body and the optic sense organ. That psychological events are intrinsically bound to neuro-muscular phenomens is not to be doubted.[14] The present study is an attempt to investigate the relation of psychological events to pharmacological responsivity.

Martin, Sackler, Sackler and Sackler[8] attempted to support a theory that endocrine dyscrasias instigate or occur concomitant with psychopathological states. A relation between histamine and various hormones of internal secretion was postulated. In animal experiments they were able to demonstrate that gonadectomy, adrenalectomy, thyroidectomy, parathyroidectomy, and thymectomy[8, 9, 10, 12] all increase the animal's ability to tolerate increasing doses of histamine. Working with groups of neurotics, and extramural and intramural psychotics, the Sacklers and van Ophuijsen[11] were able to demonstrate that separation of these groups by histamine tolerance was possible. Disintegrated individuals required larger doses of intramuscular histamine before diastolic blood pressure reached zero than did those better integrated.[2] Later, Lucy[4] was able to confirm this finding in schizophrenic and non-schizophrenic patients, and also found that long-term "deteriorated" patients tended to show higher histamine tolerances than new admissions and short-term patients.

In 1953, Steinman[13] developed a manual for the evaluation of figure drawings in which he used the criterion of histamine tolerance as an index of psychological distress, contributing a physiological measure to validation of the use of this projective technique in qualitative testing.

Problem

This study was organized to investigate the relation between histamine tolerance in patients demonstrating psychopathological states and two indices of perception; (1) visual autokinetic movement judgments, and (2) human movement responses elicited by the Rorschach Test. It was assumed that these three variables were manifestations of a common function of the organism, and hypothesized that variation in histamine tolerance, a measure of vascular reactivity, would occur concomitant with and in reciprocal relation to variation in the two measures of movement perception. Also, this set of data was relevant to the validity of the Steinman manual in providing an "estimated" histamine tolerance.

Procedure

The group of fifty-three patients, composed of 27 men and 26 women between ages of 20 to 55 attending the Clinic Division of Creedmoor Institute, was administered the Rorschach Test, the Human Figure Drawing Test, and a Visual Autokinesis Test.

The Visual Autokinesis Test was performed as follows: the subject was seated eleven feet from a dummy apparatus, painted non-reflecting black, the front of which contained a large number of pigeon holes. The stimulus was provided by a 5 watt incandescent bulb in a small portable light-proof box which had an aperture one-sixteenth of an inch in diameter, thirteen inches from the light source. The aperture was screened so as to eliminate the diffusion of light and to reduce glare. The appearance of the stimulus in a darkened room was that of a faint dot of light with negligible illumination. The subject was requested to describe the extent of the first movement of light in terms of feet and inches. Any momentary halt in the apparent movement or changes of direction terminated what was defined as the first movement. The figure used for the individual index of autokinetic movement was the mean score of five such judgments.

An index of Rorschach movement was obtained for each case by subtracting the weighted sum of the color responses from the sum of the human movement re-

[2] "Tolerance" is that dose of histamine (given intramuscularly as diphosphate and recorded in milligrams of histamine base) which reduces diastolic blood pressure to zero.

sponses (the experience balance). Human figure drawings were evaluated by means of the Steinman manual and the scores obtained were designated "estimated histamine tolerances."

Histamine tolerance ("actual", distinguished from "estimated", as later described) was obtained from a sample of twenty-four patients. Allergy, cardiac disorder and other medical states made it impossible to obtain tolerance data on the remainder of the group. In order to validate use of the Steinman manual for estimating histamine tolerance from human figure drawings, the "estimated histamine tolerances" provided by the use of the technique described in the manual were correlated with the actual tolerance of the sample group.

Results

The coefficient of correlation of .702 obtained between figure drawing and actual histamine tolerance data was demonstrated to be significant at the .004 level of confidence. However, the predictive efficiency of the correlation was only .29 above chance. This level of predictive efficiency is insufficient to allow estimated tolerance to replace actual tolerance, and the writer decided to use actual tolerance as the basic tool for analysis of data, duplicating all statistical studies with the "estimated tolerance" as a supplementary checking procedure. It was hoped that more information might be obtained on the worth of the estimated tolerances by comparing the final correlations obtained.

The M:C ratios obtained from the individual Rorschach tests were coded for statistical use and correlated with the individual estimated and individual actual histamine tolerance. The correlation coefficients −.100 and .132 respectively were definitely insignificant and from this it was concluded that there was no demonstrable relationship between the M:C ratios and histamine tolerance.

The coded M:C ratios were correlated with the individual means of autokinetic judgments, using both the sample group and the total population. These correlations −.064 and −.191 respectively were demonstrated to be insignificant and it was therefore concluded that there was no demonstrable relationship between the M:C ratios and autokinetic movement.

The estimated histamine tolerances were correlated with the individual means of autokinetic judgments, the correlation coefficient of .356 being significant at the .01 level of confidence, and it was concluded that a meaningful relationship between estimated histamine tolerance and autokinetic judgments has been demonstrated. The predictive efficiency of .065 was, however, so low as to be meaningless in terms of individual or group prediction.

The correlation between actual histamine tolerance and the means of autokinetic judgments was .765. The correlation was significant at the .004 level of confidence, and predictive efficiency was .356 above chance. This relation might be defined in terms of positive co-variance, i.e., when an individual's histamine tolerance is low, he tends to see less autokinetic movement and when it is high, he tends to see more autokinetic movement. In terms of predictive efficiency, the correlation meets the minimum requirements for group prediction. It is however, too low to allow for accurate individual prediction.

Discussion

It appears that some relationship exists between the characteristics of the phenomena of perception and physiological reactivity. On the other hand, failure to demonstrate relationship between Rorschach human movement and any of the other variables is disappointing to the hypothesis of the study. The real or subjectively experienced changes in muscular tension which supposedly accompany Rorschach movement responses have not been shown to exist in community with physical responsivity or the . . . "sheer neuro-muscular dispositions" which in part govern the production of human figure drawings.[6] While the results of this investigation show

the Steinman manual to be a relatively ineffective clinical measure, the assumption that human figure drawings are closely related to physiological reaction patterns is substantiated.

Summary and Conclusions

1. A positive relation between actual histamine tolerance and visual autokinetic movement was demonstrated which had predictive value sufficient for group data.

2. A positive relation was demonstrated between estimated histamine tolerance obtained from human figure drawings by means of the Steinman manual and actual histamine tolerance. This relationship is, however, of no predictive value for individual or group data.

3. A positive relation was demonstrated between estimated histamine tolerance and autokinetic movement. Predictive efficiency was, however, too low for any practical use.

4. A positive relation was demonstrated between actual and estimated histamine tolerance. This relationship, while significant, was too variable for predictive efficiency with either individual or group data.

5. No relation was found between (a) Rorschach human movement and visual autokinetic movement, (b) Rorschach human movement and actual histamine tolerance, or (c) between Rorschach human movement and estimated histamine tolerance.

Bibliography

1. Adams, M. F., Autokinetic sensation. *Physiological Monographs.* Chicago: University of Chicago Press, 1912.
2. Guilford, J. P., and Dallenbach, K. M. A study of the autokinetic sensation. *Amer. J. Psychol.*, 1928, 40, 83-91.
3. Kleint, H. Versuche Uber die Wahrnehmung Isolierte Bewegund. *Zeitschrift fur Psychologie.*, 1938, 142, 259-290; 1939, 143, 299-317.
4. Lucy, J. D., Histamine tolerance in schizophrenia. *Archives of Neurology and Psychiatry.*, 1954, Vol. 71, No. 5.
5. Martin, C. R., Sackler, R. R., Sackler, A. M., and Sackler, M. D. Propylthiouracil induced changes in histamine tolerance of intact and adrenalectomized rats. *Federation Proceedings*, 1953, Vol. 12, 1, 1149.
6. Murphy, G. *Personality: A Biosocial Approach to Origins and Structure.* New York: Harper and Company, 1947.
7. Rapaport, D. *Organization and Pathology of Thought.* New York: Columbia University Press, 1951, pages 15-38.
8. Rorschach, H. *Psychodiagnostics.* New York: Grune and Stratton, Inc. 1942.
9. Sackler, A. M., Sackler, M. D., Martin, C. R., and Sackler, R. R. Thymectomy and histamine tolerance. *Federation Proceedings*, 1953, Vol. 12, 1, 1196.
10. Sackler, M. D., Sackler, A. M., Martin, C. R., and Sackler, R. R., Gonadectomy and histamine tolerance. *Federation Proceedings*, 1953, 1149.
11. Sackler, R. R., Sackler, M. D., Van Ophuijsen, H. H. W., Co Tui, and Sackler, A. M. An appraisal of the current status of histamine biochemotherapy in psychiatry and the theoretical significance of its effects, Part II; Endocrinologic orientation to psychiatric disorders. *J. Clin. Experim. Psychopathology*, 1951, 12, 10.
12. Sackler, R. R. Experiments on the histamine-thyroid-parathyroid-adrenal relationship. *Federation Proceedings*, 1952, Vol. II, I, 905.
13. Steinman, K. *The Validity of a Projective Technique in the Determination of Relative Intensity in Psychosis.* Unpublished Doctoral Dissertation. New York University, 1952.
14. Woodworth, R. E. *Experimental Psychology.* New York: Henry Holt, 1938.

THE USE OF THE MMPI IN PREDICTING IRREGULAR DISCHARGE AMONG TUBERCULOSIS PATIENTS[1]

GEORGE CALDEN AND JOHN R. THURSTON
VA Hospital, Madison, Wisconsin

AND

BARBARA M. STEWART AND SHALOM E. VINEBERG
VA Hospital, San Fernando, California

Introduction

The frequency of irregular discharge from tuberculosis hospitals is a major and longstanding medical and social problem [2]. The average rate of irregular discharge from many large tuberculosis hospitals approaches 50% [6]. This indicates that for each patient who remains until he receives maximum hospital benefits, there is another patient who leaves the hospital prematurely, i.e., he receives an irregular discharge. In view of this alarming state of affairs, it appears that the modern tuberculosis therapy program cannot be applied effectively to a large proportion of tuberculosis patients.

It would be extremely advantageous to the tuberculosis hospital staff, attempting to cope with this ever-present problem, to be able to determine relatively early in hospitalization which patients would ultimately elect to leave the hospital irregularly. Once identified, these patients could be supplied with the increased attention, support, education or whatever else is necessary to convince them that a decision to leave irregularly would be unwise and detrimental to their best interests.

A promising approach to this problem seemed to be an investigation of the Minnesota Multiphasic Personality Inventory as a means of developing a psychological scale for predicting irregular discharge. Studies by Page [3], Hand [1], Rorabaugh and Guthrie [4], and Speth [5] have utilized the MMPI in investigating the personality characteristics of tuberculosis patients who receive irregular discharges. However, very little of this research has been oriented directly toward the practical problem of detecting, early in hospitalization, those patients who are likely to leave irregularly, nor has an attempt been made to integrate or reconcile the findings of these various reports.

The purpose of this paper is to present the results of an integrated research effort, to compare our findings with those of earlier studies, and to describe the development of an MMPI scale for predicting irregular discharge. This investigation was undertaken by the clinical psychology services of the VA Hospitals at Madison, Wisconsin, and San Fernando, California.

Method

Madison Study. During a period of one year, between February 1953 and February 1954, all male tuberculosis patients admitted to the VA Hospital, Madison, Wisconsin, were given the Individual Form of the MMPI as part of a routine battery of psychological tests. The MMPI was administered during the second month of their hospitalization. In all, 182 patients were tested. The irregular discharge group consisted of the first 50 patients (35 AMA, 10 AWOL & 5 DD)[2] in the sample who left the hospital without medical sanction. The regular discharge group consisted of the first 50 patients in the sample who obtained maximum hospital benefit (MHB) discharges, after having been in the hospital at least three months.

---

[1]From the Veterans Administration Hospitals, Madison, Wisconsin and San Fernando, California.

[2]There are three categories of irregular discharge: AMA—against medical advice, AWOL—absent without leave, and DD—disciplinary discharge.

San Fernando Study. The abbreviated MMPI (Group Form, 373 items) was routinely administered, usually within two weeks after admission, to male tuberculosis patients entering the VA Hospital, San Fernando, California, between September 1951 and February 1954. At the end of this period, 442 patients had been tested; 79 had received MHB discharges after at least six months of hospitalization, and 103 had received an AWOL discharge.

At each of the two hospitals, the irregular discharge group was compared with the regular discharge group on the basis of their test performance on each of the standard personality scales as well as on the individual MMPI items. The purposes of these comparisons were (*a*) to determine whether any of the existing MMPI standard scales would discriminate between groups of regular and irregular discharge patients within each hospital and between hospitals, and (*b*) if none of the standard MMPI scales were applicable, to develop a new scale based on individual items that would differentiate discharge groups both within and between hospitals.

Results and Discussion

Analysis of Standard Scale Scores. In comparing the discharge groups, it was found that the Madison irregular discharge group obtained higher T scores on most of the MMPI scales than did the Madison regular discharge group. Statistical analysis of the data indicated that the irregular discharge group obtained significantly higher scores ($p < .05$) on two scales, Depression (*D*) and Hypomania (*Ma*). A comparison of MMPI raw scores at San Fernando did not reveal a consistent elevation obtained by the AWOL patients as compared with the regular discharge patients. However, the AWOL group scored significantly higher on the Psychopathic Deviate (*Pd*) and Hypomania (*Ma*) scales ($p < .05$), and significantly lower on the *K* scale ($p < .05$).

Of the three scales which differentiated between the San Fernando groups, only one (*Ma*) was found to be significant in the Madison study; however, the direction of the Madison differences on *Pd* and *K* was the same as that found at San Fernando. With respect to these three scales, therefore, the findings at both hospitals were consistent. With few exceptions, the Madison patients obtained lower MMPI scores than the San Fernando groups. The most striking difference was found on the *D* scale, where the Madison MHB patients obtained a mean *T* score 9 points lower than either of the San Fernando groups.

The Madison and San Fernando studies were compared with two previous investigations of irregular discharge based on male veterans[1, 5]. In all four studies, significant MMPI differences were found between regular and irregular discharge groups. In three of the studies, irregular discharge patients were found to have an elevated *Pd* score, and in two a high *Ma* score was found. Differences in the *K* and *D* scales were each reported only once, and therefore cannot be considered characteristic of veteran tuberculosis patients as a whole.

Comparisons between the four veterans studies and two others based on nonveteran discharge groups are less conclusive. Rorabaugh and Guthrie[4] found no differences between their groups, nor did Page[3], when he considered male patients alone. When scores of men and women were combined, Page's irregular discharge patients obtained an elevated *D* similar to that found at Madison, but presented the enigma of a significantly *lower* score on *Pd*.

Analysis of Individual Items. Although two of the standard MMPI scales were found to differentiate male discharge groups in more than one hospital, none of the scales did so in all hospitals. The next step, therefore, was to determine the feasibility of constructing a new scale, based on individual MMPI items, that would effectively achieve this discrimination.

The results of separate item analyses conducted at the Madison and San Fernando hospitals revealed that a number of MMPI items differentiated regular from irregular discharge patients at the .05 level or better. Ninety-six of the 550 items used in the Madison study were significant. An analysis of one-half of the San Fern-

ando records (40 MHBs and 51 AWOLs) resulted in 23 significant items out of the 373 used in that study. The Rorabaugh study disclosed 48 of 550 items that were significant discriminators. However, upon inspection, it was found that only three of the significant items were common to both the Madison and San Fernando studies. Only seven of the significant items were common to both the Madison and Rorabaugh studies, while five items were found to be significant both at San Fernando and in the Rorabaugh study. Only one item (No. 224 — "My parents often objected to the kind of people I went around with" — answered *True* by the irregular discharge patients) was found to be significant in all three studies.

In addition to this cross validation study *between* hospitals, an investigation was carried out at San Fernando in order to cross-validate their 23 significant MMPI items *within* the San Fernando hospital. The MMPI records of the remaining 39 MHB and 52 AWOL patients from this hospital were utilized. Only one of the initial 23 items was found to remain significant ($p < .05$). This result would be expected on a purely chance basis. Item 224, cited above as being consistently significant among hospitals, failed to maintain its level of significance in this within hospital cross-validation.

The Irregular Discharge (Id) Scale. An Irregular Discharge (*Id*) Scale, composed of 57 significant items derived from the Madison study, was constructed in a further effort to explore this problem. (Although the Madison study yielded 96 significant items, the *Id* scale utilized only 57 of the items in order to make it applicable to other studies using the abbreviated Group Form). The items comprising the *Id* scale are listed in Table 1. The letters *T* or *F* after the items indicate the direction in which

TABLE 1. FIFTY-SEVEN MMPI ITEMS COMPRISING THE *Id* SCALE

1. 5 T	16. 76 T	31. 225 F	46. 306 F
2. 6 T	17. 93 T	32. 230 F	47. 310 F
3. 9 F	18. 96 F	33. 234 T	48. 312 T
4. 19 T	19. 100 T	34. 238 T	49. 316 T
5. 25 T	20. 101 T	35. 239 T	50. 317 T
6. 27 T	21. 102 T	36. 242 F	51. 330 F
7. 28 T	22. 118 T	37. 243 F	52. 338 T
8. 32 T	23. 122 F	38. 247 T	53. 344 T
9. 33 T	24. 124 T	39. 251 T	54. 347 F
10. 36 F	25. 145 T	40. 253 F	55. 364 T
11. 41 T	26. 167 T	41. 272 F	56. 398 T
12. 43 T	27. 186 T	42. 280 T	57. 406 T
13. 46 F	28. 213 T	43. 291 T	
14. 55 F	29. 223 T	44. 294 F	
15. 61 T	30. 224 T	45. 303 T	

irregular discharge patients tend to respond in contrast to regular discharge patients. The numbers of the items correspond to those used in the Group Form. An irregular discharge score is obtained by adding the total number of statements, in an individual record, that are answered in the direction shown in the table.

When applied to the original criterion groups, i.e., the Madison sample, the scale effectively discriminated between the types of discharge (50 regular and 50 irregular discharge patients) at a high level of significance ($p < .001$). The mean *Id* score of the 50 regular discharge patients was 14.92 and the mean score of the 50 irregular discharge patients was 25.18.

Examination of the score distributions, as shown in Table 2, revealed that the scale was maximally effective if a cutting score of 18 was employed to distinguish discharge groups. A score of 18 or above was suggestive of irregular discharge. Using this cutting score, it was possible to correctly identify the discharge status of the patients in the Madison sample in 75% of the cases.

The *Id* scale was applied to the San Fernando discharge groups (79 MHBs and 103 AWOLS) in order to determine its effectiveness on another sample. The mean score of the MHB group was 16.72 and that of the AWOL group was 19.88. This

TABLE 2. DISTRIBUTIONS OF *Id* SCALE SCORES FOR DISCHARGE GROUPS AT MADISON AND SAN FERNANDO

Scores	Madison		San Fernando	
	MHB	Irregular	MHB	AWOL
42-45		1		
38-41		2	1	
34-37		2		2
30-33		9	1	4
26-29		13	5	15
22-25	5	5	13	21
18-21	10	8	12	25
14-17	15	7	19	19
10-13	14	3	19	11
6-9	3		7	5
2-5	3		2	1
Total	50	50	79	103
Mean	14.92	25.18	16.72	19.88

difference in means is significant at the .01 level. Using the cutting score of 18, derived from the Madison study, the scale would have correctly predicted the discharge status of the San Fernando sample in 63% of the cases. A Chi-square analysis of this set of predictions showed significance at the .01 level of confidence. Although the accuracy of prediction has diminished from 75% in the Madison study to 63% in the San Fernando application, the *Id* scale consistently demonstrates statistical significance in identifying irregular discharge patients at both hospitals.

From the above results, it appears that the *Id* scale may prove useful both as a research technique and as a practical device for the detection of irregular discharge risks.

SUMMARY

This paper describes an integrated research effort aimed at obtaining an MMPI scale for predicting irregular discharges among tuberculosis patients. None of the standard MMPI scales consistently discriminated between regular and irregular discharge patients. However, an Irregular Discharge (*Id*) Scale, consisting of 57 MMPI items, effectively discriminated discharge groups at two of the hospitals.

REFERENCES

1. HAND, T. J. Personality characteristics of a tuberculosis group. *Am. J. Phys. Med.*, 1952, 31, 95-101.
2. LEWIS, W. C., LORENZ, T. H. & CALDEN, G. Irregular discharge among tuberculous patients. *Psychosom. Med.*, May, 1955.
3. PAGE, R. B. Application of the Minnesota Multiphasic Personality Inventory to tuberculosis patients. Unpublished doctoral dissertation, University of Minnesota, 1949.
4. RORABAUGH, M. E. and Guthrie, G. The personality characteristics of tuberculous patients who leave the tuberculosis hospital against medical advice. *Am. Rev. Tuberc.* 1953, 67, 432-439.
5. SPETH, E. W. An investigation of emotional and social concomitants in the irregular discharge rates of hospitalized patients with tuberculosis. Unpublished doctoral dissertation, University of Pittsburgh, 1953.
6. TOLLEN, W. B. Irregular patients' discharge: the problem of hospitalization of the tuberculous. Veterans Administration pamphlet, 10-27, 1948.

A COMPARISON OF RIGHT HEMIPLEGICS WITH LEFT HEMIPLEGICS IN VISUAL PERCEPTION*

NANCY E. WOOD

Cleveland Hearing and Speech Center

Problem

The theory of *localization* versus *non-localization* of function has become a complicated academic argument resulting in extensive research on cerebral dominance from various areas of professional interest. Since the diagnosis of hemiplegia presumes that one hemisphere is damaged and the other is not, a portion of that research has been done with hemiplegic populations. Some writers[8, 9] have reported that right hemiplegics (damage to the left hemisphere of the brain) show more devastation to speech, language, intelligence and other highly integrated mental activities than left hemiplegics (damage to the right hemisphere of the brain). Others[1, 7] have reported a qualitative rather than a quantitative relationship between the side of the lesion and intellectual functioning.

Since visual perception requires integrated mental activities and an analysis of visual symbols, it becomes necessary to study visual perception and intelligence concordantly. It is not the purpose of this paper to define or discuss visual perception, nor is it the purpose to establish the specific tests used as tests of visual perception. This study is concerned with whether or not right hemiplegics and left hemiplegics respond differently when visual stimuli are presented in a structured procedure.

Method

The subjects in this study consisted of 50 right and 36 left spastic hemiplegics within an age range of 10 to 30 years, with no history of institutionalization. They had been cleared psychiatrically and were known to be free from any extreme emotional involvements. Each subject had been medically diagnosed as hemiplegic without complications of other cerebral palsy symptomatology. All selected subjects were assumed to be genetically predisposed right-handed individuals since there was no history of familial left handedness. The subjects were examined for aphasia[4], hearing loss and/or severe speech problems before acceptance for further testing.

An evaluation of information obtained for each subject concerning time of lesion, chronological age, race, sex, seizures, and socio-economic level revealed no significant differences between the groups. Each subject was examined for extinction phenomena[2], motor involvement[11, 14], and intelligence[12,14], and the two groups were found to be comparable when these data were analyzed.

Prior to the use of the selected visual perception tests, the visual acuity of each subject was evaluated[6] with respect to simultaneous vision, vertical imbalance, lateral imbalance, fusion, right eye acuity, and steriopsis. Of the right hemiplegics 34% were found to have good vision as compared with 47% of the left hemiplegics. Both groups revealed more disturbance in lateral balance, right and left eye acuity, and/or steriopsis, and both groups responded more accurately to tests of simultaneous vision, vertical imbalance and/or fusion. The general trends observed in visual acuity were similar for the two groups, but for purposes of clarity, the data obtained for the groups with good vision and the groups with poor vision were treated separately.

Materials and Results

The selected visual perception tests included three measurements: figure-ground relationships, perceptual closure, and visual perseveration. The procedure for all

*Nancy E. Wood (Ph.D. Northwestern University, 1952) is Assistant Director of The Cleveland Hearing and Speech Center and Assistant Professor in the Graduate School, Western Reserve University. This paper is abstracted from a dissertation prepared under the direction of Dr. Harold Westlake, Head, Department of Speech Correction and Audiology, Northwestern University and the research committee composed of Drs. Carhart, Gaeth, Lightfoot, Moore, and Myklebust. The writer is particularly grateful to Dr. Meyer A. Perlstein, who acted as Medical Director for this study and to the National Society for Crippled Children and Adults, specifically the Michigan Chapter of that Society, for securing subjects, medical histories and financial support for this project.

tests was the same. The test stimuli were exposed tachistoscopically in rooms that could be darkened and where no light or other distractions were present. The subject was seated in an ordinary chair and the projection screen was placed directly in front of him at a distance of 12 feet. Movement was not permitted during testing so that this standard distance could be maintained. Standard conditions were strictly adhered to during the entire testing procedure, and all images were projected in the same manner and at the same exposure (½ second). All test pictures were drawn to scale, so that the images were the same size and, when projected, occupied a space of about one foot square on the screen. The entire procedure was explained before the testing began. After all of the slides were presented and the responses recorded, the stimulus pictures were presented again for an indefinite period of time. All of the subjects were able to correctly identify all of the slides when they were exposed for this longer time interval, supporting the contention that the objects were not unfamiliar items.

Figure-ground Relationships. The test selected to evaluate figure-ground relationships was patterned after a technique originated by Werner and Strauss[13] and used by others[3] with brain injured children. The test consisted of nine figures representing simple line drawings of everyday objects imbedded in structured homogeneous backgrounds. Each picture was presented on a separate slide and the subject described what he had seen after the image had been tachistoscopically exposed.

Each subject was found to respond to each simulus picture in one of five ways: (A) response to the ground but no mention of the figure, (B) response to the ground and incorrect or obscure mention of the figure, (C) correct response to both figure and ground, (D) no mention of the ground and incorrect or obscure response to the figure, and (E) no mention of the ground but correct response to the figure. An individual response trend was determined for each subject. If seven of the nine responses were types (A) or (B), the subject was classified as being *predominantly ground* in response. If the subject responded to seven of the nine stimulus pictures as discussed in (C), (D), and (E), the subject was said to be *predominantly figure* in response. When these data were treated the differences in responses between the right hemiplegics and the left hemiplegics were negligible regardless of visual acuity.

It was possible that a difference might be determined between the groups, if the subject's response to each individual stimulus picture was considered separately. Therefore, each response to each stimulus picture was analyzed and if the subject responded to the figure only, whether this response was accurate or not, the response was classified as a *figure response.* Any mention of the ground was tabulated as a *ground response.* The resulting data for the good vision groups and the poor vision groups were treated separately. None of the obtained data approached the necessary requirements for significance, revealing negligible differences between the two groups regardless of visual acuity. This suggested that although the right hemiplegics of both the good and poor vision groups showed a slight trend toward more ground responses than the left hemiplegics, this trend could not be considered significant.

Perceptual Closure. The test selected for measuring perceptual closure was devised by Street[10]. Originally, some of the pictures in this test consisted of black lines on white backgrounds and others consisted of white lines on black backgrounds. In order to reduce the *figure-ground* component and stress the characteristic of *closure,* all of the pictures were adapted to black lines on white backgrounds. The slides consisted of 12 images of well known objects. The subject was asked to describe what he saw and the responses were recorded as being either correct or incorrect. The identification of the pictures required the subjects to perceive an object as a *whole,* even though only a portion of the object was visible.

The obtained Chi-squares became zero in seven of the twelve trials in the good vision groups; in five of the twelve trials in the poor vision groups. None of the resulting data were significant at a 5% level of confidence. Apparently the right and left hemiplegics were not significantly different in their responses to these perceptual closure stimuli.

Visual Perseveration. An adaptation of an experiment by Harrower[5] was used to determine the perseverative trends of the two groups. The basic slide was an ambiguous vase-profile (after Rubin) so structured that normal response to this image fluctuates constantly between a response to the vase and a response to the profile. There were three slides in which the vase-aspect was reinforced in continuing degrees and three slides in which the profile-aspect was reinforced in similar manner. The basic slide was projected on the screen and the subject was told that this slide could represent either a vase or two profiles. If the subject identified the basic slide as a vase, the next slide presented was the first reinforced profile image. A reversed procedure was used if the subject identified the basic slide as two profiles.

Perseveration was indicated in 27% of the right hemiplegics as compared with 31% of the left hemiplegics. When the groups were divided with respect to visual acuity and the data were treated, no difference was found between the two groups with good vision since, after correction for continuity, the Chi square became zero. In the poor vision groups, 19% of the right and 37% of the left hemiplegics showed indication of visual perseveration. Although there appeared to be almost twice the proportionate perseverative responses in the left hemiplegics with poor vision as compared with the right hemiplegics with poor vision, the differences were not significant at the 5% level of confidence.

It is necessary to report that there were no significant differences between the right and the left hemiplegics with respect to the three measurements of visual perception studied here. However, the slight trend suggesting that the left hemiplegics with poor vision showed a higher proportion of perseverative responses than the right hemiplegics with poor vision, might suggest further exploration.

Summary

The purpose of this report was to compare a group of right spastic hemiplegics with a group of left spastic hemiplegics to determine if differences in responses to structured visual stimuli occurred between the two groups. Both groups showed evidence of disturbance in responses to selected measurements of figure-ground, perceptual closure, and visual perseveration regardless of visual acuity but the differences between the responses of the right hemiplegics and the left hemiplegics were negligible. Although this population of hemiplegics showed evidence of sensory involvement in addition to the impairment of motor functions, a positive relationship between the side of the brain lesion and a disturbance in visual perception cannot be supported by this study.

References

1. Anderson, A. L. The effect of laterality localization of focal brain lesions on the Wechsler-Bellevue subtests. *J. clin. Psychol.*, 1951, 7, 149-153.
2. Bender, M. B., Shapiro, M. F., and Schappell, A. W. Extinction phenomena in hemiplegia. *Arch. Neurol. Psychiat.*, 1949, 62, 717-724.
3. Dolphin, J. E. and Cruickshank, Wm. M. The figure-background relationship in children with cerebral palsy. *J. clin. Psychol.*, 1951, 7, 228-231.
4. Eisenson, J. *Examining for Aphasia.* New York: Psychological Corp., 1946.
5. Harrower, H. R. Changes in figure-ground perception of patients with cortical lesions. *Brit. J. Psychol.*, 1939, 30, 47-51.
6. *Keystone Telebinocular.* Meadville, Penna.: Keystone View Company.
7. Lipscomb, P. R., and Krusen, F. H. Cerebral Palsy with special reference to physical therapy. *Arch. Phys. Ther.*, 1943, 24, 342-349.
8. McIntire, J. T. A study of the distribution of physical handicap and mental diagnosis in cerebral palsied children. *Amer. J. ment. Def.*, 1947, 51, 624-626.
9. Phelps, W. M. Motor handicaps and retardation. *Proc. Amer. Assoc. ment. Def.*, 1938, 43, 26-31.
10. Street, R. F. A gestalt completion test. Columbia University: *N. Y. Bur. Pub.*, 1931, 11-65.
11. Van Der Lugt, M. J. A. *Psychomotor Scale for Measurement of Manual Ability.* Beverly Hills, California: Western Psychological Serv., 1948.
12. Wechsler, D. *The Wechsler-Bellevue Intelligence Scale, Form II.* New York: Psychological Corp., 1946.
13. Werner, H, and Strauss, A. A. Pathology of figure-background relation in the child. *J. abnorm. soc. Psychol.*, 1941, 36, 236.
14. Wood, N. E. A comparison of right hemiplegics with left hemiplegics in motor skills and intelligence. (Publication pending.)

THE PATIENT SPEAKS: PSYCHOLOGIST-PATIENT RELATIONS[1]

ALVIN R. HOWARD

Veterans Administration Hospital, Chillicothe, Ohio[2]

What follows is abstracted from a recorded interview between a former mental patient (S) and the author (E_1). A psychiatrist (E_2) also participated toward the latter part of the interview.[3] The occasion was S's visit, after an eight-year interval, to the institution where he had previously been confined for four months. Several weeks preceding this latter visit he had stopped in to renew acquaintances and, on that occasion, had been invited by the Chief, Professional Services, to return. During each visit he addressed a group of staff members.

It is felt that this report will aid the psychologist in evaluating his role in the field of interpersonal relations. Frequently we have opportunities to infer a patient's feelings, but less often the chance to hear an *ex*-patient expound. In this instance, S also wields the not unimportant aegis of social acceptability as a result of his current achievements in the business world. Thus it is an approved member of society who recalls his experiences and presents his ideas. This writer does not suggest the immediate practice of all these ideas; actually, some of them would soon disintegrate under the intense heat of critical scrutiny. He does feel, however, that S has provided scarce cross-validation data in the field of psychologist-patient relations.

Background Data

It may prove helpful to review briefly some of the events in S's background.[4] He is the oldest of four sibs and came to the United States from South America at the age of eight. Uninterested in formal schooling, he left when in the tenth grade. Following two years in military service, he was discharged as psychoneurotic. In the meantime, he had married a woman three years his senior.

Shortly after discharge from the Army, he completed high school. Soon afterward he was hospitalized with the diagnosis: "Schizophrenia with compulsive urges to violence, including suicide." He was placed on a "Special Observation" ward and a subsequent staff decision was rendered against removing him from that milieu. However, one week later, a letter from his wife was received requesting that S "be treated intelligently if any cooperation is expected of him." As a result of her apparent deep concern for his welfare and depth of understanding of his condition, the staff removed him from the "Special Observation" ward within one week following receipt of her letter. Within four months, S had undergone electroshock and insulin therapies and was granted a trial visit home. Later, the wife again wrote the hospital:

> "S has been steadily progressing in the past year. Only once did I find him faltering, but not enough to cause serious worry about him. He has been trying to adjust himself, both at home and with the outside world. I feel he has done miraculously well. True it is a slow process which cannot be forced along, but the makings of progress and effort on his part are there."

Since his discharge from the hospital, S has remained away from institutions. He has four children and owns his home. He is Director of Personnel for a small manufacturing concern and thus has daily contacts of an interpersonal nature.

Regardless of the motives that may be ascribed to his re-visiting the hospital where he was formerly an involuntary patient, S does offer diverse comments of

---

[1]From the Veterans Administration Hospital, Chillicothe, Ohio.

[2]Now Chief, Clinical Psychology Service, Veterans Administration Hospital, Sheridan Wyoming.

[3]Dr. Henry Luidens, then Chief, Professional Services, Veterans Administration Hospital, Chillicothe, Ohio. Now Manager, Veterans Administration Hospital, Coatesville, Pennsylvania.

[4]Data have been changed to prevent identification of S.

interest to all psychologists. These are now presented in the hope that they will lead to increased understanding of the interplay between the psychologist and his patient.

S feels quite strongly that the psychologist initiates his evaluation much too rapidly. He prefers that the inquiry be delayed so as to secure subsequently greater cooperation:

S: Well, one thing, your psychologist's testing generally occurs when the patient is brand new here, and being brand new here, regardless if he has been in a previous hospital or just a first admission, he's all excited, he's tense, he tries to readjust to the – uh - environment here and out of the clear blue sky you pop in, you start asking questions. Now he doesn't know but that you're gonna be the guy that's supposed to be his salvation here. He's going to want to latch on to you. Or on the other hand, you are the last guy he sees; the doctor sees him first and everybody else sees him and then you come along. By this time his patience has worn thin; he's been questioned, he's been prodded, he's been needled, he's been checked and double-checked, and then you start with this bunch of gibberish, or at least what seems to him to be gibberish, and his patience is worn pretty thin. Now, like I said on the lawn, supposing you were to go into the ward first, even while the Psychia- uh - delay your interview for a week or so, but go on the ward periodically and get him—get him accustomed to seeing you and to—uh—seeing you on the ward and gain some confidence, then when he comes in here, he'll think that it was through a special dispensation on your part that he came here and he'll be very cooperative. It might—he might envision future field trips to come out here.

S insists that the patient be approached as an adult. He resents the "pedantic" and "pompous" attitude of the psychologist and again cites his opposition to hasty "probing".

S: That's one point. Another, I wish psychologists would drop that pedantic atmosphere, that business of "Now this is just a simple little test, let's make a game of it. Let's see what you can do". The patient that's out of contact, he doesn't need no preamble. He's going to be uncooperative anyhow, and the patient that is in contact, if you treat him like an adult—uh—he'll respond as an adult. So—uh . . .

E_1: I take it that you don't go along with this notion that the mental patient now is similar to a child.

S: I'll go further. I've got to go back to my parakeet. My parakeet doesn't know a human being or what it might do to him, and has a natural fear. Now, the patient and re-re-repeti—uh—repetition—uh through kindness and association—you can get that patient to respond to you. The bird, I keep putting my finger in, now he perches on it. By perching, I put some bird seed on my mouth and now he's learned to pick it off my mouth. He finds he gets food there. I put birdseed on my shoulder. I gently put him there. He flies away. I pick him up again, put him back, he finally settles. It's the same thing with a patient. I don't care how disturbed, or how out of contact the patient is, often he'll associate you with something. You can see here in the institution itself, you'll find patients who are out of contact, yet when the attendant asks him to go to chow, he falls in line and goes with them. Now, it isn't that the hunger—that he's associated hunger and the word "chow"—but he has also associated the pa-, the attendant. The attendant is the source of the chow, so the attendant is a good guy. Uh- the psychologist—uh—I hate to say it, too often is pompous. Really and truly, every psychologist I've ever seen—he's going to find a sore spot. But by the time—in a hospital especially such as this, the sore spot's already been lost. That's what you're probing for now. And it's not going to be done in a week or a month because remember most of the patients that have come through here have already been screened through receiving hospitals and intermediate hospitals and finally they've just been abandoned here, and you're getting under the very rough end of it here. So I think instead of jumping on them the first week or so, just go in and just make contact visits with them, just find out how responsive he is or unresponsive he is on the ward before you drag him in here. At least that's my attitude on it. I—uh—frankly, with no offense to you or your profession, I have no use for psychologists. Because to me—they are always asking me—uh—what to me are insulting and degrading, well filthy questions. Perhaps I came in contact with the wrong psychologists. That may be.

Once again S refers to the desirability of "social" visiting prior to a more formal evaluation. He also proposes ways of utilizing the layman's confusion of psychologist and psychiatrist for the patient's advantage, but his prime thesis remains his belief in an extended "warm-up" period.

E_1: Well, in general, S, you said something about psychologists in general, that they're anathema to you and — uh—what suggestions would you offer now, that is over and above what you've already said, to any psychologist in terms of how he may facilitate—

S: Better contact—

E_1: Yeah—get along better with the patients.

S: Well, there is a little evil there which is the psychologist has only a brief amount of time with the patient and then he is permanently through with him. The initial testing and then you

don't see the psychologist at all. And, as a result, I think both the psychologist and the patient don't really care whether they warm up to each other or don't warm up, because—"What the hell, after this interview, we're through. You're not going to see me anymore. You can't give me a pass. You can't do me any good." That's why I suggested this pre-visit on the ward. Uh—getting out on the ward when the patient comes in and showing yourself to him. Maybe giving him a cigarette. Don't - don't get involved with psychology - just simply a social visit. You've come to see a prospective customer of yours. Then later on when he comes here he'll be warmer—he'll want to cooperate with you. He'll recall your kindnesses or, at least, he will recall that he's seen you before. This way, I don't know if you're my judge here, or whether you're (supposed?) to assign me to a different ward or what it is I don't know, and you don't tell. You just sit there and you want me to look at some pictures. Now, why? Well, if I'm lucid you're gonna say it's to help the doctor to determine my difficulty. "But look, Mac, I have no difficulty. I just want out. I just drink too much. That's why I'm in here."

E_1: Well S, tell me this. You said earlier this afternoon when we were walking over to 3 building that the—the patient doesn't care too much about the psychologist unless he can think of him as his doctor. If he knows now, as I pointed out to you, we tell the patient that we are not his doctor, that we funnel, as the social worker, and the contact man and the occupational therapist—

S: They're going to see that.

E_1: Well, if now, the patient doesn't see that and doesn't care to speak with anyone except his psychiatrist, his ward physician, then what difference does it make if the psychologist makes these social visits, before he tests the man?

S: Because at that particular point, that patient is still new enough here where he doesn't realize that you can't do him any good in the way of passes, or ward changes, or special privileges. So you make a couple of visits—it takes a month or so before you're fully conditioned to the hospital—to realize that you're a lost cause. You won't help him one damn bit, that is, personally help him.

E_1: Yeah.

S: Aside from the team operation. In other words what I'm conspiring with you to do is to trick him into confiding in you while the iron's hot. Because once he cools off, you'll never get him in here to give you responses.

E_1: What do you do then? Here's a man, let's say, that comes in this afternoon and we are both psychologists and we walk over to Building 31. Now what do we do?

S: Go out into the ward. You've been doing this, making one visit a day. Some of the patients that were there yesterday will see you again today. Some of your prospective customers, you'll be seeing them twice. Some of them you see once. In other words, you'd be a regular fixture in their minds. You come in there, you have the key privilege, which means that you're a somebody there. The attendants call you "sir" or "doctor". In fact, use your title, on the wards, there especially, misleading as it may be later on, at that particular moment it's darned good psychology, and that way they think you're part of a team of doctors that eventually is going to work on them. Now, they don't know which doctor is going to work on them at this point in 31. They're just there for observation. And it is from there that you draw your patients for here, isn't it? So you see, literally for the moment, you are their doctor. So if you keep making a daily visit, the new patient comes in, some get transferred out but—uh—as long as you keep making one visit a day there, they're gonna get to see you. Just go around—"How are you?" "How do you feel?" and—look, you put a little honey out there, maybe you'll catch a fly.

E_1: Well, now you walk over there, and you tell him "I'm Dr. E_1," and you chat with him socially and then, as you say, some of the people that you met yesterday are there today. It's the second social visit for them. Now, what happens, S, when the man says to you, "Yes, Dr. E., but what do you do here?"

S: O. K. 99% of all the patients you have here are just as ignorant as I was. The word psychologist and psychiatrist mean the same thing. "I'm a psychologist. I'm a clinical psychologist." The word "Clinic" is well known. You're a clinical psychologist. Now, do I have to know any more? Let them—Don't oversell yourself. Just—that's all you have to sell.

E_1: What do you mean when you say "don't oversell yourself?"

S: Because what you can wind up doing—you can wind up making them think you're a psychiatrist, which is not my intention. You're the clinical psychologist. You're all these boys' doctor. You're nobody's special doctor. You see, you're a staff doctor. That's O. K. Later on they'll get their own individual doctor. At this point, they don't know yet they're going to have their own doctor on the ward, that there's a ward doctor taking care of a group of patients. Up to that point, he's still responsive to you. The rest of your staff, by the same token, now they haven't got the title to use. They come out there—"I'm the psychologist". Period. Don't say any more than that. Don't make no promise of any kind. Uh—As soon as you say you're a psychologist—for some reason it's like to my—my—uh—son seven years old—"Well, who makes it rain?" — "God" — period. That takes care of everything. I don't have to get involved with nimbus or cumulus or low pressure areas or high pressure areas. "God" — period. "Psychologist" —period. That's it. Who would dream of questioning any further than that?

E_1: We do have some patients, of course, who don't know what the term "psychologist" means—

S: They wouldn't know what a psychiatrist was either.

E_1: What happens when these people ask you though? What do you do?

S: All right. "I assist your ward doctor in helping to find out what your sickness is." "But Doc—I ain't sick." O. K. That's for you to find an answer to that, because there's very few of them here who will admit to being sick. And that's a big problem—to get them to admit it. Now —I can't solve a lot of these problems—again—I'm just uh—

E_1: Certainly not, S.

S: I'm just suggesting—I do think though that the pre-visit before testing would do a world of good.

E_1: Sort of a warm-up period with the patient, is that it?

S: Yeah. Not necessarily sitting down with him, but just circulating around in the ward, making a point of saying, "How are you?", "What's your home town?", "Are you married?". In fact, you can find out a lot of preliminary information just by questioning around right here on the ward, much more readily than you will here. Because either he'll sit here and become an exhibitionist or he'll clam up tighter than a drum.

E_1: Well, what is there about an office like this that gets such a variety of reactions?

S: The word "office" in itself. It's an office and it's not the ward . . .

Forced to play a role in the center of the therapeutic stage, S requests an explanation of the events surrounding him. One-way communication antagonizes the patient whose self-perception is that of a person who can "*think* for himself." Here we are reminded both of his wife's request that he "be treated inteligently if any cooperation is expected of him" and also of S's own statements regarding an "adult" approach to the patient.

S: Uh, nobody ever took the time to explain *anything* to me whether it was electric shock or homosexuality or—or any other terminology used, *nobody* wanted to explain it to me. In essence the attitude taken was, "It's none of your business, we'll take care of everything. You just answer the questions." And I felt, and I think justifiably so, it's my life. I'm certainly entitled to some consideration. Now this is in contradiction to what I just told you a few minutes ago. Uh, whose right is it over my life—I mean I'm here *non-compos mentis*—you have a right to do what you want with me. True. But I said when a patient cannot *think* for himself—but if I am sitting here and I ask you to tell me what it means, am I committing a crime? I would think I am entitled to an answer. I mean I'm the subject of the thing. Certainly you'll—you'll treat me anyhow—even if you don't tell me and—uh—if I resist you it will make the job a little more difficult, but you will still manage to get the information you want one way or another. But the thing is, this makes me feel better, and as a psychologist you should certainly want to make me feel better to make—to calm me down to accept this hospitalization—and I think you would be doing me a service by—therapeutic service by—explaining it to me.

S briefly reiterates his dislike of "filthy" questions and then indicates having picked up a smattering of psychology as the sole result of "pumping" from psychiatrists and psychologists.

S: Did you love your father more than your mother? Did you ever want to "screw" your mother? Some psychiatrists and some psychologists have no respect for sensibility. Now I know all the dirty words necessary to—for the description of the anatomy of both male and female but, on the other hand I also know the proper words for them too. Yet some psychologists and psychiatrists insist on using the —uh—filthy—uh—uh—that is the street jargon for it. Now there again, find out first the background of the patient. If he can talk intelligent English then for "God's sake don't refer to—uh—uh—to his penis as a "pecker", call it a "penis". It won't hurt you.

E_2: Yeah.

S: It's—it's real name happens to be a penis.

E_2: Mm hmm.

S: On the other hand if the guy wouldn't know what a penis is then call it a—uh—in the language that he can understand. But uh—uh—there again is the idea of respecting the dignity of the—of being a human being. Because, believe it or not, outside of a dirty joke, I don't like to discuss, for example, the physical anatomy of my wife with any man.

E_2: No.

S: Nor do I desire to identify my physical anatomy with *you*. Uh—to you. There again—Dr. Freud—that Freudian homosexuality. I'm - I'm getting a homosexual contact that way. I don't want it.

E_2: How much Freud have you read incidentally?

S: None.

E_2: You picked this all up from psychiatrists and—

S: Psychiatrists and psychologists—

E_2: Yeah.

S: Pumping away at me.

Although these comments by a former resident of a mental hospital may not be representative (in a strictly statistical sense) of patients-at-large, yet he does afford us some insights into areas that require intensive review by psychologists and related professional persons. He re-arouses questions in us that remain to be answered by a continuing research effort. Will observation of our patient in informal settings contribute to more effective understanding of him? What is the most effective time relationship between "informal observation" and "more formal evaluation?" Are we emphasizing immediate "service needs" of an institution to the detriment of understanding our patients sufficiently? What measures need be taken to educate the public about the necessity for our functioning on the therapeutic team? What are the psychologist's functions in orienting the patient to the institutional milieu? These and related questions must be resolved if we are to enhance our relationships with patients and other persons concerned with their welfare.

SUMMARY

The recorded comments of a former mental hospital patient are presented as a means of increasing the psychologist's understanding of his role in various interpersonal relationships. The subject expresses his attitude on various issues of direct concern to psychologists and offers suggestions for improving the relations between psychologist and patient. His views reinforce the need for more critical evaluation of various functions of the psychologist.

SOCIOMETRIC CHOICE AND THE SZONDI TEST[1]

DAVID PEARL AND DURAND JACOBS

Veterans Administration Hospital, Battle Creek, Michigan

Veterans Administration Hospital, Marion, Indiana

PROBLEM

The Szondi test, according to the theorizing of both Szondi [4] and Deri [2] may be viewed as a test of interpersonal relations. According to their rationale, factors determining Szondi test behavior of subjects would also underlie the choice-rejection aspects of their interpersonal relations. It would follow then, that subjects with similar Szondi test behavior would demonstrate similarity in their interpersonal evaluations. Conversely, one would expect those with similar affective judgments of people to be similar in Szondi test choices. The present study was directed at investigating this inference. It was one of several which studied various aspects of sociopsychological dynamics within a privilege ward at a Veterans Administration neuropsychiatric hospital. Hypotheses formulated were:

1. Subjects with similar affective reactions to their wardmates, as indicated by sociometric analysis, would agree among themselves a significantly greater number of times in their evaluations of Szondi test photographs than would subjects with contradictory or unrelated sociometric choice-reject patterns.

2. Subjects with similar sociometric patterns would have significantly fewer disagreements in Szondi test choices than those with contradictory or unrelated ward choice patterns.

METHOD

Thirty-nine adult male veteran patients assigned to a privilege ward served as subjects. These patients formed a heterogeneous diagnostic group, including schizophrenic, neurotic and characterological diagnoses. All subjects had been ward members for at least two weeks and had ample opportunity to develop affective attitudes

[1]From the Veterans Administration Hospital, Battle Creek, Michigan.

toward each other. In individual sessions, each subject was initially asked to rate the photographs of his wardmates on a continuum of liking, according to Q technique procedure[3]. This was followed by the administration of the Szondi test.

By intercorrelating the sociometric choice patterns of these subjects, it became possible to determine those with similar, contradictory, and unrelated affective evaluations of the ward population. A choice pattern correlation of .41, being significant at the P .01 level was adopted as the criterion of similarity. Forty pairs of subjects with correlations meeting this criterion were selected as constituting those most similar in their inter-personal evaluations. For comparative purposes, a second group consisting of the 40 subject pairs with the largest negative correlations was formed. These two groups represented the extremes of sociometric agreement and any relationship of Szondi to sociometric choice similarity should be demonstrated most clearly by this comparison.

A previous analysis of the sociometric intercorrelations for all 39 subjects had revealed four clusters of patients. The members of each cluster tended to show significantly similar choice patterns. Conversely, with the exception of cluster 4, their choice patterns tended to show no significant relationship to that of other patients not in their cluster. These clusters may be viewed as representing the types of affective choice patterns within the ward structure. As another test of the hypotheses, it would be expected that cluster members' Szondi choices should be significantly more similar than those not falling within such a grouping. To permit such comparison, six patients were selected as a control group from the remaining non-clustered subjects. The average intercorrelation for these latter was .07 as contrasted to .40 for the cluster subjects.

For these tests of the hypotheses, comparisons were made with respect to the number of (a) agreements for Szondi photos rated as liked, (b) agreements for those rated as disliked, (c) total disagreements (a photograph liked by one and disliked by the other member of the two subjects being compared), and (d) total agreements.

Two approaches were employed in the analysis of this data. In the first, the number of agreements or disagreements were to be viewed as scores, and analyses of variance and comparisons of means were to be utilized. In the second, these numbers were to be treated as freqencies and chi squares were to be computed.

Results

Data and results of the comparison of Szondi photograph choices for positively and negatively correlated sociometric pairs are given in Table 1. Analyses of differences between means, utilizing *t*-tests, disclosed significant differences for three of the four comparisons, namely, total agreements, agreements for disliked photos, and number of disagreements. A chi square analysis of data provided similar results for these comparisons.

Table 1. Comparison of Szondi Choices for 40 Positive and 40 Negatively Correlated Sociometric Choice Pairs

Category	Positively Correlated Pairs	Negatively Correlated Pairs	Mean of Positively Correlated Pairs	Mean of Negatively Correlated Pairs	*t*	Confidence level
Agreements for "liked" photographs	164	151	4.10	3.78	1.03	P.30
Agreements for "disliked" photographs	205	164	5.13	4.10	2.59	P.01
Total number of agreements	369	315	9.23	7.88	2.26	P.05
Number of disagreements	145	184	3.63	4.60	2.26	P.05

Analyses of variance were made for the four sociometric groups and their control subjects for each of the comparisons listed above. Means of the various groups for each choice category and the results of the analyses of variance are given in Table 2. F ratios, significant at the P .05 level were obtained in three comparisons. In these categories, the various sociometric groups were compared to the control subjects and the results are summarized in Table 3.

TABLE 2. ANALYSES OF VARIANCE OF SZONDI CHOICES FOR SOCIOMETRIC AND CONTROL GROUPS

Category	Means					F	Confidence Level
	Group 1	Group 2	Group 3	Group 4	Control Subjects		
Agreements on "liked" photographs	3.7	4.1	5.7	3.3	3.3	4.062	Not signif.
Agreements on "disliked" photographs	5.8	5.3	4.0	3.7	2.9	5.893	P.05
Total number of agreements	9.5	9.4	9.7	7.0	6.2	8.427	P.05
Number of disagreements	3.2	3.3	3.8	3.7	6.1	7.803	P.05

Except for two comparisons involving group 4, and one of group 3, highly significant differences were found between groups and the control subjects. Since the group clusters were homogeneous with respect to sociometric agreement, one would expect that no differences in cluster internal consistency would be found when the sociometric groups were compared between themselves. This expectation was borne out, as no significant differences were found when the in-group agreements were computed.

TABLE 3. COMPARISON OF SOCIOMETRIC GROUPS AND CONTROLS FOR SZONDI CHOICES

Category	Control Subjects vs. Group 1		Group 2		Group 3		Group 4	
	t	Confidence Level	t	Confidence Level	t	Confidence Level	t	Confidence Level
Agreements for "disliked" photographs	3.766	P.001	4.138	P.001	1.390	P.20	.728	P.50
Total number of agreements	3.537	P.01	4.539	P.001	3.719	P.001	.650	P.60
Number of disagreements	3.963	P.001	4.964	P.001	3.078	P.001	2.510	P.02

The failure to obtain significant differences between group 4 and the other groups may be due to the marginal quality of this group. It is composed of only three subjects whose average intercorrelation is smallest of all the sociometric clusters. Furthermore, as previously pointed out, these subjects had many more significant extra-cluster correlations than did members of the other sociometric groups. For these reasons, this group was dropped from consideration in further analyses.

Chi square analyses between the groups for the number of total like and dislike agreements, as well as number of disagreements, likewise revealed significant differences (P .001). As in the t-test analyses, these were differences between the control subjects and the sociometric groupings.

Discussion

These findings support hypotheses 1 and 2, which were inferred from certain aspects of Szondi theory. It must be stressed, however, that the Szondi genetic framework is only one of many which would encompass these results. Borstelmann and Klopfer[1] have pointed out the many serious shortcomings in this framework. Further examination of the Szondi data fails also to lend support to the Szondi diagnostic framework, in that internal contradictions are found. Since significant differences in the degree of consensus for Szondi test stimuli were found between the homogeneous and control sociometric groups, we would expect to find significant differences in one or more categories in a similar analysis of consensus for the eight Szondi diagnostic categories. However, no significant differences were found in agreements or disagreements for any of these categories. Deri[2] has also hypothesized that the frequency of selection within a given category is a valid measure within the need system represented. The subsequent expectation that the extent of reaction to various categories would differ significantly among the homogeneous and control sociometric groups was not borne out by the data. This would suggest that either Deri's hypothesis is gratuitous or that the breakdown of the stimuli photographs into these particular diagnostic categories is unwarranted.

A consideration of these results suggests that whatever merit the Szondi test possesses is due to the employment of the photographic medium. By providing cues, the photographs permit subjects to project their interpersonal attitudes and feelings. The Szondi and Deri mediating formulations are thus not essential. One may hypothesize that similarities in affective attitudinal patterns toward the Szondi test photographs and the photographs of wardmates are due to underlying personality processes. Individuals having such personality processes in common would tend to react similarly in affective tone to the cues provided by photographs or individuals with whom they have contact. Data to be reported in detail in a subsequent publication suggest that such processes are related to the self concept. Ratings of the applicability of self concept statements given by members of the sociometric and control groups, when categorized and treated by analysis of variance, disclosed significant interaction between sociometric groups and self concept categories.

Summary and Conclusions

Forty pairs of patients with significantly similar ward sociometric choice patterns were compared in Szondi test choices to forty ward patient pairs with the most dissimilar sociometric choices. In a like manner, four clusters of patients, each similar in the sociometric patterning of its members, were compared to a control group of patients whose choice patterns were unrelated. Results support hypotheses affirming that significant differences would be found. While consistent with Szondi test theory, these results do not necessarily support such theory. Indeed, further analysis of the data discloses internal contradictions in the Szondi diagnostic framework. The results are seen as most satisfactorily explained within the framework of general perceptual theory.

References

1. Borstelmann, L. J. and Klopfer, W. G. The Szondi test: a review and critical evaluation. *Psychol. Bull.*, 1953, *50*, 112-132.
2. Deri, S. *Introduction to the Szondi test, theory and practice.* New York: Grune and Stratton, 1949.
3. Stephenson, W. *The Study of Behavior: Q-technique and its methodology.* Chicago: University of Chicago Press, 1954.
4. Szondi, L. *Triebdiagnostick.* Bern: Hans Huber, 1947.

TEETH-GRINDING: SOME RELATIONSHIPS TO ANXIETY, HOSTILITY, AND HYPERACTIVITY*

FRANCIS F. VERNALLIS

Veterans Administration, Pittsburgh, Pa.

Problem

The habit of teeth-grinding is approached in this study as a problem in personality structure. The emphasis is on the relationship between a behavioral characteristic and three personality factors rather than the onset or development of the characteristic. Many dentists recognize teeth-grinding as a source of dental pathology. Miller [2] recognizes the teeth-grinding habit as a dental syndrome which he has designated as "bruxism." Bruxism is defined by Miller as, "Excessive gritting of the teeth from nervousness, nerve irritations or lesions." Other dental authorities corroborate Miller's definition of the habit, and two dental nomenclature bodies have given it official recognition. But some dental investigators retain reservations concerning the psychological etiology of teeth-grinding owing to the possibility of its possessing a predominantly systemic or mechanical origin or even of its being a natural activity.

Psychoanalytic speculation concerning teeth-grinding has regarded the habit as a resultant of fixation at the oral-sadistic stage of psychosexual development. Only the described behavioral characteristics of the oral-sadistic personality were used by the writer for evolving testable psychological variables. Psychoanalytic theorems as to the etiology of teethgrinding were not probed by this study. In the general area of nervous habits, divergent experimental results have been reported in the psychological literature. Olson [3] conducted one of the earliest studies in this sphere and he demonstrated that the subject was amenable to objective inquiry. He discovered a positive correlation between items of a personality inventory and oral nervous habits. Studies which have endeavored to establish nervous habits as reliable symptoms of psychopathology have been unsuccessful.

Method

The subjects were 40 teeth-grinders and their controls drawn from the student body of The Pennsylvania State University. They were individually matched for sex and intelligence, and group matched for age and educational status.

The identification of teeth-grinders constituted a major problem. As a partial solution, a questionnaire was constructed which combined items known to be symptomatic of the habit and direct questions. Respondents were requested to state directly whether they knew they gritted their teeth and whether another person had called attention to it. The questionnaire was administered to approximately 950 students and it succeeded in identifying a statistically adequate sample. Individual interrogations for evidence of the habit's presence were carried out before the subjects were admitted into the testing group. The subjects gave numerous verbal reports which had an authentic ring to them as evidence for teeth-gritting. For example, one subject stated, "I know I grind my teeth because I tap out the harmony with them when I listen to music." Another, "I have a recurring dream of my teeth crumbling and when I awaken my teeth are firmly clamped together. When I'm awake I feel like I'm chewing my teeth out of my mouth." And a third, "My girl friend has noticed my habit and she is openly annoyed by it." Teethgrinding proves to be a source of considerable distress to those given to the habit. Control subjects were selected on the basis of their possessing the matching characteristics and freedom from the habit. Several of them displayed other overt nervous mannerisms.

The test battery consisted of the Rorschach Content Test devised by Elizur [1], four Blacky Pictures, the Manifest Anxiety Scale by Taylor [4], the cooperativeness-

*Submitted in partial fulfillment of the requirements for the degree of doctor of philosophy at The Pennsylvania State University, 1953.

hostility scale of the Guilford-Zimmerman Temperament Survey, and the MMPI hypomania scale. In addition, the writer formulated eight items which were also intended to measure hostility. The test battery was administered individually to each subject.

Results

Biserial correlation was the statistical method used for analyzing the scores of the Rorschach Content Test, the Manifest Anxiety Scale, and the cooperativeness-hostility and hypomania scales. The four Blacky Pictures and the eight hostility items devised by the writer were analyzed by means of chi square. The results are given in table 1. The anxiety measures revealed statistically significant relationships

Table 1. The Comparison of Teeth-Grinders and Their Controls on Three Personality Factors

Factor and Instrument	r-bis	X^2	Sig. Level
Anxiety			
RCT-anxiety	.29		5%
Manifest Anxiety Test	.36		1%
Hyperactivity			
MMPI Hypomania Scale	.32		1%
Hostility			
RCT-hostility	.57		1%
G-Z hostility scale	.22		10%
Blacky Pictures			
oral-sadism		7.86	5%
castration-anxiety		.80	n.s.
guilt-feelings		2.80	10%
narcissistic-love-object		.46	n.s.
Hostility items		*	n.s.

*Chi square analysis disclosed only one of the eight items was significant.

with teeth-grinding. Divergent results were found with respect to hostility. The hostility scale of the Rorschach Content Test yielded the highest correlation in the test battery, a .57 biserial *r* significant beyond the one per cent level. Card II, oral-sadism, of the Blacky Pictures also gave confirmatory evidence for the hypothesis; its probability level was beyond one per cent. The other three Blacky cards are only tangentially related to hostility so their non-significant probabilities are viewed as test artifacts to the hostility: teeth-grinding hypothesis. The latter cards would have been omitted had not the scoring procedures of the Blacky required additional cards. The Guilford-Zimmerman hostility scale revealed a statistically insignificant relationship although it was in the predicted direction (ten per cent level). The writer's hostility items were even less effective since merely one of the items—Are sarcasm and cutting wit your favorite forms of humor?—reached the five per cent level. A biserial *r* of .32 was found between the MMPI hypomania scale and teeth-grinding.

Discussion

The acceptance of the hypothesis that teeth-grinding and anxiety are related is plausible, for it conforms to dental and lay opinion which holds that "worry" and "nervousness" go together with the habit. The role of hostility in bruxism is less obvious but would be suspected to be present by personality theorists. Although the test results are somewhat contradictory on this dimension, the projective instruments do sustain a relationship and the Guilford-Zimmerman hostility scale suggests a positive association. The negative results obtained with the writer's hostility items are probably due to poor item construction. The correlation between hyperactivity and the criterion appears to be a rather commonsense one. Thus, this study provides psychometric evidence for widely held notions concerning the teeth-grinding habit.

As for the practical application of these findings, they suggest teeth-grinding is a reliable clue to underlying tension, although not necessarily serious psychopathology. When the symptom appears on mental examination, for example, the inference of the presence of anxiety, hostility, and hyper-activity is indicated. Since psychological conflicts underly the condition, the study also suggests treatment should be aimed at these factors for its correction.

Summary

Teeth-grinding is recognized among dentists as a pathological dental syndrome designated as "bruxism." There is some dispute as to whether its etiology is psychological or systemic and mechanical. The subjects were 40 teeth-grinders and their controls drawn from the student body of The Pennsylvania State University. Teeth-grinders were identified by means of a questionnaire and personal interview. Psychometric examination suggests a positive relationship between bruxism and the personality factors given in the title.

References

1. Elizur, A. Content analysis of the Rorschach with regard to anxiety and hostility. *J. Proj. Tech.*, 1949, *13*, 247-284.
2. Miller, S. C. *Textbook of Periodontia.* Philadelphia: Blakiston Co., 1950.
3. Olson, W. C. *The Measurement of Nervous Habits in Normal Children.* Mineapolis: Univ. Minn. Press, 1929.
4. Taylor, Janet A. The relationship of anxiety to the conditioned eyelid response. *J. Exp. Psychol.*, 1951, *41*, 81-92.

A STUDY OF IDEAL PSYCHOLOGICAL ADJUSTMENT

ANDERS SWEETLAND AND GEORGE FRANK

Florida State University

Problem

Psychotherapists are usually aware of the predicted course and goal of the individual case, but overall goals for the practice of psychotherapy, couched in terms of the ideally adjusted individual, have not been formulated. If it were possible to obtain an adequate description of the so-called "well-adjusted" personality, the therapist might have a more objective goal. Although many studies have elicited differences between the maladjusted and the normal personality, no study has attempted to determine the characteristics of the opposite end of the continuum. The purpose of this study was to make a first approximation of "ideal psychological adjustment." Essentially, we are looking for some reasonably concrete information to use as a point of departure for further research.

Method

The procedure was a replication of that used by Stephenson[6], in testing the validity of the Jungian extraversion-introversion hypothesis. A large sample (approximately 1000 items) of traits of adjustment were culled from the literature. As far as possible these were taken verbatim. One hundred of these were selected randomly, to be Q-sorted as follows:

Envelope 1: These (2) are most like me of all.
Envelope 2: These (14) are much like me, but less than those in envelope 1.

Envelope 3: These (34) are a little like me, less than those in envelopes 1 and 2, but more like me than the rest.
Envelope 4: These (34) are slightly unlike me.
Envelope 5: These (14) are much unlike me.
Envelope 6: These (2) are least like me of all.

The selection of the subjects was based on scores on the Minnesota Personality Scale (MPS). This Scale is a factor-analytically derived test of personality, constructed by using the most reliable test items from the Minnesota Scale for the Survey of Opinions, the Minnesota Inventory of Social Attitudes, and Bell's Adjustment Inventory. Item analysis was performed and those items which did not consistently yield a critical ratio of above 3.0 were deleted, making the internal consistency high. The final scale permits a judgment of items on the basis of a five-point scale. Odd-even reliability coefficients are consistently above .90.

The scale consists of five subscales representing five factors obtained from the analysis of the intercorrelations of each item with each other item. These are: Morale (attitudes towards the legal system, to education, and general adjustment); Social Adjustment (feelings of inferiority, social behavior, social preferences, etc.); Family Relations (parent-child adjustment); Emotionality (health and emotional adjustment). The fifth scale, Economic Conservatism, was not used in this study as it was not felt to be pertinent.

The subjects were selected on the basis of percentile scores. Scores between the 35th and 70th percentiles on the four scales of the MPS were considered to be evidence of average adjustment, and those above the 75th percentile were designated as superior adjustment.

Two groups of twenty subjects were selected from the university population, matched for sex and college level. Q-sorts from the forty subjects were intercorrelated by the Pearson product-moment method, and the matrix factor analyzed by Thurstone's centroid method [7].

We hypothesize that if the criterion of selection was a valid one, the groups should sort out into two fairly distinct groups on the Q-sort. Should this be the case, it would then be reasonable to assume that those variables which were most characteristic of the group with MPS scores above the 75th percentile were, likewise, characteristic of the ideally adjusted or of ideal adjustment. Similarly, those personality characteristics of the group with MPS scores between the 35th and 70th percentiles, were characteristic of average adjustment.

Results[1]

Factor analysis isolated two separate factors. Factor I, common to both the average and superior groups, appears to be a factor of general adjustment. It is characterized by such statements as:

I am able to love.
I enjoy people's company.
I realize that I am not perfect.
I get along well with my brothers and sisters.
I get along well with my parents.

and to exclude such traits as:

I do not try to do more than I am able.
I can avoid situations which produce anxiety in me.
I dream at night.

Factor II separated the subjects into two distinct groups: those individuals who were high in this factor, and those who were low. Interestingly, in terms of the aforementioned hypothesis concerning the criterion of selection (see Method), the Q-sort

[1]The actual matrix of intercorrelations of subjects and the table of factor loadings may be obtained from the ADI where they have been deposited.

isolated people in this factor in terms of the initial selection of subjects. Those subjects who had high MPS scores in general sorted out together (17 out of the 20), and those subjects who had average MPS scores were in general grouped (13 out of the 20). Since those with the high MPS scores were also the ones with the high factor loadings in this factor, it was assumed that we had isolated the characteristics of the ideally adjusted in this factor. The factor is characterized by such statements as:

I am able to love.
I enjoy living.
I am not overly dependent upon my family.
I have the ability to be part of a group and yet able to differ from that group when necessary.
I am self-dependent.
I can accept love.
I have the ability to think in terms of a group, society, the world, and not only in terms of myself.
I realize that I not only have good points, but bad ones also.
I am tolerant, that is, I can understand another's beliefs, practices, and habits, without necessarily sharing or accepting them.

and these individuals consistently reject such traits as:

I have the ability to form strong and lasting emotional ties.
I avoid becoming over-tired.
I do not try to do more than I am able.
I am sentimental.
I have escaped from the conviction of sin.
I have some tendency towards jealousy.
I am idealistic.
I am patient.
I have a desire to be a member of some group, either a club, society, a culture, etc.
I dream at night.

Discussion

Factor I, common to both groups, and apparently characteristic of adjustment in general, appears to be characterized by a rapport with self and environment. In this respect it is not unlike the findings of research concerned with the individual following psychotherapy as reviewed by Rogers[5]. Interestingly enough, the adjusted personality does not feel himself to be free from anxiety.

Factor II, having isolated the subjects in terms of the criterion of selection scores on the MPS, appeared to be unique with the superior group. This factor was also characterized by rapport with self and environment, but there are some interesting differences. Both groups say, "I am able to love," but the superior group adds, "I can accept love." Both indicate the presence of anxiety, but the anxiety in the superior group appears to be more realistic and focalized. Further, the superior group tends to reject the "Jack Armstrong, All American Boy" concept of adjustment, by denying such traits as: interest in physical activity, sentimentality, idealism, or even patience.

Although both groups feel that they get along well with parents and sibs, the superior group adds: "I am not overly dependent upon my family;" "I am self dependent." This is similar to what Angyal[1] and Maslow[4] have called "autonomy." In terms of research in psychotherapy, Lipskin[3] and Rogers[5] found that clients, following therapy, perceived themselves as being more independent.

The average group tended to delimit their group identification to the family, while the superior group identified with larger social units: "group, society, the world." The superior group, however, also says, "I have the ability to be part of the group, and yet able to differ from that group when necessary," and, "I am tolerant, that is, I can understand another's beliefs, practices, and habits, without necessarily sharing or accepting them." Maslow[4] has referred to this as a quality of detachment, a freedom from the immobilizing effect of emotional involvement with others. The superior group seems to crystalize this with the statement, "I (do not) have the ability to form strong and lasting emotional ties."

In another phase of this research, Greenberger[2] attempted to obtain a scale of "Ideal Adjustment" on the Minnesota Multiphasic Personality Inventory (MMPI). As he used the same subjects, it was possible to make some comparisons between the Q-sort data and the MMPI which he found differentiated the average from the superior groups. It was found that in using the Q-sort, the subjects tended to place items pertaining to aberrant behavior toward the middle of the continuum, neither markedly like or unlike. Greenberger's results, on the other hand, showed that items

of pathology were rated as "false," i.e., unlike, on the MMPI. The discrepancy in results appears to be a function of the refined scaling technique utilized in the Q-sort data, as compared with the bi-polar attitude set up by the MMPI, and suggests that in actuality, very well adjusted people have little need to actively deny aberrant behavior as part of their personality.

Summary

1. An attempt was made to obtain a first approximation of the nature of ideal psychological adjustment. Twenty average and twenty well-adjusted college students, selected on the basis of their scores on the Minnesota Personality Scale, Q-sorted 100 personality traits descriptive of adjustment. Factor analysis of the data yielded two factors.

2. Factor I, common to all subjects, appeared to be characterized by rapport with self and environment. It also included some feelings of anxiety.

3. Factor II, unique to the superior group, showed a similar, but not identical, pattern of rapport with self and environment. In addition, the superior group rejected, as unlike them, the qualities of idealism, patience, and desire for physical activity. The superior group also described themselves as being relatively detached and independent in their relationships with people and social groups, and showed little need to actively deny the presence of pathological behavior in their personalities.

4. The results leave us not with a scale, but with a first approximation of ideal adjustment composed of 19 traits. We still recognize the need to investigate the rest of the 900 items thus far untapped, in order to develop an adequate scale of ideal adjustment, as well as the need to test the reliability of these findings.

References

1. Angyal, A. *Foundations for a science of personality.* New York: Commonwealth, 1941.
2. Greenberger, H. A tentative scale of good-adjustment for the MMPI. Unpublished master's thesis, Fla. State Univer., 1953.
3. Lipskin, S. The client evaluates nondirective psychotherapy. *J. consult. Psychol.*, 1948, *12*, 137-146.
4. Maslow, A. Self-actualized people: A study of psychological health. *Pers.*, 1950, Symposium #1, 11-34.
5. Rogers, C. R. *Client-centered therapy.* New York: Houghton-Mifflin, 1951, Ch. 4.
6. Stephenson, W. A methodological consideration of Jung's typology. *J. ment. Sci.*, 1939, *85*, 185-205.
7. Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univer. of Chicago Press, 1947.

GOAL SETTING BEHAVIOR, EXPRESSED FEELINGS OF ADEQUACY, AND THE CORRESPONDENCE BETWEEN THE PERCEIVED AND IDEAL SELF

LOUIS T. LEPINE AND BERNARD CHODORKOFF

Catholic University of America

Veterans Administration Hospital, Dearborn, Michigan

PROBLEM

Cohen [3] recently investigated goal setting behavior within a level-of-aspiration framework by studying two variables which were presumed to be related to the implicit factors determining goal level settings. These two variables were feelings of adequacy and self-acceptance. It was found that goal level setting was not related to feelings of adequacy, but that a curvilinear relationship did exist between goal level setting and self-acceptance.

The present study was undertaken to see whether modification of the procedures used by Cohen would lead to clearer results. It was felt that measures of adequacy feelings and self-acceptance as obtained by judges' ratings based on Rorschach protocols might not be as meaningful as similar measures derived directly from the subjects (Ss) themselves. Also, it was felt that a measure based on a level of aspiration task may be more meaningful, if the task subjects each S to similar experiences, and if each S has to face a comparison between his performance and the performance of his peers.

The level of aspiration task used was a code-deciphering procedure involving pre-arranged performance scores. Correspondence between the perceived and ideal self was taken as a measure related to and paralleling self-acceptance. A measure of expressed feelings of adequacy was obtained from the way each S sorted self-descriptive statements. The methodology to be described had the advantage of enabling us not only to explore further Cohen's work, but also to continue investigating the relationship of correspondence between the perceived and ideal self to other variables. The following hypotheses were tested:

1. The more closely related to actual performance the individual's goal settings are, the greater will be the correspondence between his perceived and ideal self.

2. The more closely related to actual performance the individual's goal settings are, the greater will be his expressed feelings of adequacy.

3. The greater the correspondence between the individual's perceived and ideal self, the greater the degree to which he will express feelings of adequacy.

METHOD

Subjects. The subjects were 30 hospitalized veterans who were between the ages of 20 - 40 years. Although half the group consisted of NP patients and half of medical and surgical patients, these were combined and treated as one group because there were no significant differences between these types of patients on any of the variables considered here.

Procedure. Each S was seen twice. During the first session S was asked to sort 125 self-descriptive statements into a forced distribution so that he described the sort of person he saw himself to be (perceived self).[1] Following this, he was given a level of aspiration procedure. This consisted of 20 one-minute trials of deciphering letters into their code counterparts. At the end of each trial, the S was given a score which supposedly reflected how well he did in comparison to other patients like himself. Before beginning the next trial, the patient was asked to tell how well he actually

[1]For a more complete account and description of this Q-sort see [3].

expected to do on the next trial. Prearranged scores were used so that each S was given the same sequence of performance scores. In this way, the situation was set up so that each S responded to the same overall performance conditions.

Although the overall trend was in an upward direction, the prearranged scores sometimes went up and sometimes went down. Because of this situation some of the trials could be considered as representing improving performance conditions. The average prearranged performance score under improving conditions was 55, and under poorer performance conditions was 52, so that poorer or improving performance was always a relative thing.

The second session was held the following day. S was again asked to sort the self-descriptive statements, but this time describing the person he would like to be (ideal self).

Analyses. Various scores were derived:

1. *Correspondence score.* S's sorting of his perceived self was correlated with the sorting of his ideal self. The product-moment r obtained was then transformed to Fisher's z.

2. *Goal discrepancy score.* This score was the mean difference between performance and subsequent aspiration bid for all test trials.

3. *Goal discrepancy score for relatively improving performance.* This score was the mean difference between performance and subsequent aspiration bid for all trials on which S's level of performance was improving.

4. *Goal discrepancy score for relatively poorer performance.* This score was similar to the above one with the exception that it was derived from trials on which S's level of performance was becoming relatively poorer. Scores 3 and 4 were included because it seemed likely that our Ss would respond differentially to improving or poorer performance conditions. Inasmuch as performance scores were prearranged, specific trials could be selected at which time performance was relatively improving or becoming relatively poorer.

5. With the aid of five judges, items from the 125 self-descriptive statements were selected because they reflected feelings of adequacy or inadequacy. On the basis of S's sorting of these items, an *expressed feelings of adequacy score* was derived.

Results and Discussion

The intercorrelations between all variables were plotted; there were no suggestions of curvilinearity so product-moment r's were computed. The significant correlations found were:

1. The *goal discrepancy score under improving performance* correlated significantly with the *goal discrepancy score* ($r = .98$).

2. The *goal discrepancy score under poorer performance* correlated significantly with the *goal discrepancy score* ($r = .98$) and with the *goal discrepancy score under improving condition* ($r = .93$).

3. The *adequacy score* correlated significantly with the *correspondence score* ($r = .70$).

These results indicate that the goal discrepancy scores were highly intercorrelated and did not show significant relationships to either expressed feelings of adequacy or to correspondence between perceived and ideal self. Hypotheses 1 and 2 were not supported. It has been assumed that correspondence between the perceived and ideal self was related to self-acceptance. Theory suggests this to be so[4]. These results, if our assumption is correct, do not agree with Cohen's findings of a curvilinear relationship between goal level settings and self-acceptance. In agreement with Cohen's findings are our results which show no significant relationship between goal level settings and expressed feelings of adequacy. Methodological

differences between the two studies make it difficult to further compare or assess these results.

In the process of trying to understand the results obtained, it appeared that the consistency of the Ss' goal setting behavior, under the changing performance conditions, might be related to expressed feelings of adequacy and the correspondence between the perceived and ideal self. This relationship might exist even though goal setting behavior *per se* was not similarly related. A study by Brownfain [1] investigating the stability of the self-concept suggested an interesting hypothesis. Brownfain suggested that the individual lacking in self-esteem was likely to be "situation-dominated". His behavior would depend more on how his environment evaluates him, rather than on his self-evaluations. In the present study we would expect the goal setting behavior of Ss who express feelings of adequacy to a greater degree, to be less dependent upon the changing performance scores given them, than would be the goal setting behavior of individuals who express feelings of adequacy to a lesser degree. As a corollary, we would expect greater correspondence between the perceived and ideal self to also be characteristic of Ss whose goal settings are less dependent upon the changing performance scores given them.

To test these hypotheses, the expressed feelings of adequacy and correspondence scores were correlated with two sets of difference scores. The first difference score was the difference between the mean goal discrepancy score for improving performance trials and the mean goal discrepancy score for poorer performance trials. The second difference score was the difference between the mean goal level (or aspiration level) for improving performance trials and the mean goal level for poorer performance trials. It was found that expressed feelings of adequacy correlated significantly with the goal discrepancy difference score ($r = .43$) and with the goal level difference score ($r = .46$). Correspondence between the perceived and ideal self did not correlate significantly with either difference score ($r = .18$, $r = .20$ respectively). These results support the hypothesis suggested by Brownfain. Goal setting behavior under improving and poorer performance conditions varied more, the more the S expressed feelings of adequacy, while the goal setting behavior of those Ss who felt less adequate were more dependent upon past performance and showed a more consistent relationship to past performance. The corollary hypothesis was not supported.

Hypothesis 3 was supported by these results. It was found that the more adequate an individual felt, the greater was the correspondence between his perceived and ideal self.

Summary

An attempt was made to study the interrelationship between goal setting behavior, expressed feelings of adequacy, and the correspondence between the perceived and ideal self. It was found that the more an individual tended to express feelings of adequacy, (1) the greater was the correspondence between his perceived and ideal self, and (2) the less dependent upon environmental evaluation of his past performance was his goal setting behavior, when environmental evaluation indicated sequential change in adequacy of performance.

Reference

1. Brownfain, J. J. Stability of the self-concept as a dimension of personality. *J. abnorm. soc. Psychol.*, 1952, *47*, 597-606.
2. Chodorkoff, B. Self-perception, perceptual defense, and adjustment. *J. abnorm. soc. Psychol.*, 1954, *49*, 508-512.
3. Cohen, L. D. Level of aspiration behavior and feelings of adequacy and self-acceptance. *J. abnorm. soc. Psychol.*, 1954, *49*, 84-86.
4. Rogers, C. R. *Client-centered therapy*. New York: Houghton Mifflin, 1951.

A COMPARISON BETWEEN THE PEACE TIME PSYCHIATRIC CASUALTY RATES OF PARACHUTISTS AND NON-PARACHUTISTS*

CHARLES WINDLE AND HAROLD E. PARKER, MAJOR, M. C.
HRU No. 3, Ft. Benning, Ga. *82d Airborne Division*

Introduction

Airborne outfits have the reputation of being elite, and paratroopers have become stereotyped as tough, daring, and aggressive soldiers[1, 10]. The difference between Airborne and non-Airborne soldiers springs primarily from the selection involved in becoming Airborne. Soldiers who volunteer for Airborne training have been found more trouble prone, more attracted by adventure and danger, more self-confident, and more attracted by discipline than most soldiers[2]. Further selectivity occurs from screening and attrition during training. Men who are not in good physical condition, who are afraid of heights, who are neurotic, or who do not believe that Airborne outfits are better than most outfits are most likely to fail[2, 11, 12, 15].

Just as Airborne soldiers differ from non-Airborne soldiers in background, personality, and attitudes, their duty as elite troops is more demanding, both in war and peace time. During war they have been called upon to perform critical missions in which mobility and surprise were crucial[4]; in peace time Airborne outfits are the most impressive show troops. Always they must perform the fearful task of parachute jumping.

Whether Airborne selection and training adequately prepare soldiers psychologically for the anxiety-arousing task of parachuting can be measured in the psychiatric casualty rate. Thatcher[14] has suggested that the emotional imbalance produced temporarily by parachute jumping might lead to the development of neuroses. Kepecs[9] has elaborated this suggestion by presenting case histories of neurotic reactions in parachutists. Further, the unique stresses of jumping might be expected to produce different types of psychiatric breakdowns than appear in most soldiers. This has been found true of other uniquely stressful situations[5, 6, 7, 13]. Grinker and Spiegel[7, p. 104-105] claimed that as a result of the paratroop group attitude of repressing fear "neurotic reactions among the paratroopers are apt to take the form of conversion symptoms involving the lower extremities." Bernstein[1] has gone further to suggest that "a significant degree of immaturity reaction is probably common to all Airborne soldiers." The purpose of this study is to determine the extent and manner in which parachutists differ from non-jumpers in peace time psychiatric breakdowns.

Method

To obtain an estimate of whether paratroopers seem able to cope with the stresses of jumping, apart from combat, a record was obtained of all jumper personnel from the 82d Airborne Division receiving psychiatric attention during the period 1 May 1953 - 15 April 1954. Recognizing that variations in psychiatric diagnoses occur among military stations[3], a control group of non-jumpers was obtained under the same referral and diagnosis conditions. This control group consisted of non-jumpers assigned to the 82d Airborne Division; similar psychiatric data were collected for this population. Since jumpers and non-jumpers were fairly well intermingled within the division and were diagnosed by the same two psychiatrists generally following SR-40-1025-1[16], there was little possibility of systematic differences in referral or diagnostic policies between jumpers and non-jumpers.

Results

(a) *The overall psychiatric casualty rate* of 82d Division jumpers and non-jumpers is shown in Table 1. Proportionately fewer jumpers received psychiatric attention than did non-jumpers. A difference as great as the observed one would be expected to occur less than one time in a thousand on a chance basis.

---

*The research reported here was conducted by the authors while the senior author was employed by The George Washington University, Human Resources Research Office, operating under contract to the Department of the Army. Opinions and conclusions are those of the writers and do not represent views of the University or the Department of the Army.

TABLE 1. PSYCHIATRIC CASUALTY RATE OF 82D DIVISION JUMPERS AND NON-JUMPERS FROM 1 MAY 1953 TO 15 APRIL 1954

Factors	Jumpers	Non-Jumpers	Total
Mean Enlisted Assigned Strength	11,942	3,302	15,244
Number of men seen by division psychiatrist	295	193	488
Estimated percent of men seen by psychiatrist (in 11½ months)	2.47%	5.84%	3.20%
	$X^2 = 95.07$; $P < .001$		

(b) *Diagnostic categories* were next compared. The frequency with which jumper and non-jumper personnel were assigned each of the most frequent diagnostic labels is shown in Table 2. Since many men exhibited more than one symptom, these

TABLE 2. NUMBER OF 82D DIVISION JUMPERS AND NON-JUMPERS DIAGNOSED AS SHOWING VARIOUS PSYCHIATRIC SYMPTOMS

Diagnostic Category	Number of Jumpers	Number of Non-Jumpers	Jumper Non-Jumper Ratio	Probability of Ratio** Being Same As: Division Jumper: Non-Jumper enlisted strength	Jumper: Non-Jumper population seen by psychiatrist
82d Airborne Division mean enlisted strength	11,942	3,302	3.6:1		
Men seen by Division Psychiatrist	295	193	1.5:1	<.001	
Dissociative Reaction	10	2	5.0:1	NS*	<.20*
Passive-Aggressive Reaction	18	5	3.6:1	NS	<.10
Amoral & Anti-Social Personality	21	6	3.5:1	NS	<.10
Alcoholism & Drug Addiction	28	9	3.1:1	NS	<.05
Behavior Disorders (including mainly Chronic AWOLism)	28	10	2.8:1	NS	<.10
Immaturity Reaction (type unspecified)	13	5	2.6:1	NS*	<.30
Anxiety	36	18	2.0:1	<.05	NS
Acute Maladjustment Reaction	8	4	2.0:1	NS*	NS*
Sex Deviates (excluding homosexuality)	9	5	1.8:1	NS*	NS
Symptomatic Habit Reactions	7	4	1.8:1	NS*	NS*
Post Head Injury Syndrome	25	15	1.7:1	<.02	NS
Conversion Reaction	31	25	1.2:1	<.001	NS
Depression	11	10	1.1:1	<.01*	NS
Headaches	31	29	1.1:1	<.001	<.20
Inadequate Personality	13	12	1.1:1	<.01	NS
Nervous System Diseases	19	17	1.1:1	<.001	NS
Somatization Reaction	7	12	.6:1	<.001*	<.05
Schizoid Personality	6	10	.6:1	<.001*	<.10
Passive-Dependency	5	9	.6:1	<.001*	<.10
Homosexuality	3	12	.3:1	<.001*	<.01

*Chi-square tests in which Yates' correction was employed because of expected cell frequencies of less than 5. Since this investigation is exploratory, direct probabilities have not been calculated [3].

**NS is used to indicate that these results could happen by chance more than 30 times in 100.

diagnostic categories cannot be considered independent. Comparisons were made between the proportion of jumpers and non-jumpers in each diagnostic category, based on both the mean enlisted jumper:non-jumper ratio in the strength of the division, and the jumper:non-jumper ratio in the cases receiving psychiatric attention.

Non-jumpers predominated in almost all diagnostic categories. Jumpers predominated in only two categories: dissociative reactions and passive-aggressive reactions. Within the psychiatric population, however, jumpers tend to show alcoholism, passive-aggressive reactions, behavior disorders and amoral or anti-social personality most frequently. Non-jumpers show more homosexuality, passive-dependency, schizoid personality and somatization reactions.

Summary

The suggestions that parachutists suffer psychiatric disorders because of the personality characteristics which induce them to volunteer for this unique type of service or because of the hazardous nature of their duty have not been borne out by the present data. Apparently these suggestions have sprung from examination of individual cases manifesting situationally-oriented symptoms without adequate recognition of control data. Jumpers have a lower psychiatric casualty rate than non-jumpers, despite their more stressful duty. This unexpected finding suggests that selection for and screening during Airborne training eliminates a high proportion of psychiatrically unfit soldiers from the successful jumper population.

On the other hand, the observations that paratroopers are prone to aggressive types of reactions[10] seems to be correct when a contrast is made against other types of breakdowns. Although there appears to be no difference in the proportion of jumpers and non-jumpers with passive-aggressive reactions, amoral or anti-social personality, alcoholism or behavior disorders, if a jumper *does* receive psychiatric attention he is more apt to fall into one of these categories than is a non-jumper psychiatric casualty.

Bibliography

1. Bernstein, R. Mental reactions of the airborne soldier. *U. S. Armed Forces Med. J.*, 1950, *1*, 1301-1306.
2. Cogan, E.A., *et al. A study of airborne volunteers: I. A comparison between volunteers for the airborne and other basic trainees (non-volunteers). II. A comparison between volunteers who successfully complete airborne training and those who fail.* The George Washington University, Human Resources Research Office, Washington, D. C., January 1953.
3. Edwards, A. L. On "the use and misuse of the chi-square test"— The case of the 2 x 2 contingency table. *Psychol. Bull.*, 1950, *47*, 341-346.
4. Gavin, J. M. Cavalry, and I don't mean horses. *Harper's Mag.*, April 1954, 54-60.
5. Gramlich, F. W. A psychological study of stress in service. *J. Gen. Psychol.*, 1949, *41*, 273-296.
6. Great Britain Air Ministry. *Psychological disorders in flying personnel of the Royal Air Force investigated during the war, 1939-1945.* London: His Majesty's Stationery Office, 1947, 344 pages.
7. Grinker, R. R. and Spiegel, J. P. *Men under stress.* York, Pa.: Maple Press Company, 1945.
8. Hunt, W. A. and Wittson, C. L. Some sources of error in the neuropsychiatric statistics of World War II. *J. Clin. Psychol.*, 1949, *5*, 350-358.
9. Kepecs, J. G. Neurotic reactions in parachutists. *Psychoanal. Quart.*, 1944, *13*, 273-299.
10. Neel, S. H. Medical aspects of military parachuting. *Milit. Surg.*, 1951, *108*, 91-105.
11. Ross, W. D., Dancey, T. E. and Brown, F. T. Rorschach scores of parachute troopers in training. *Bull. Canad. Psychol. Assn.*, 1943, *3*, 26-27 (not seen; refer to *Psychol. Abs. 17*:3231).
12. Satter, G. A. *An evaluation of the personal inventory for predicting success in parachute school.* O. S. R. D. Report 4870, 1945; Publ. Bd. No. 13293. Washington, D. C.: Dept. of Commerce, 1946.
13. Star, S. A. Psychoneurotic symptoms in the Army. In S. A. Stouffer *et al.*, (Eds.) *Studies in social psychology in World War II. Vol. II. The American soldier: Combat and its aftermath.* Princeton: Princeton University Press, 1949, pp. 411-455.
14. Thatcher, J. O. Paratroops and para-physicians. *Nav. Med. Bull.* (*Wash.*), 1942, *40*, 280-281.
15. Walk, R. D. Self-ratings of fear as a research instrument in fear-invoking situations. Unpublished paper read at the Eastern Psychological Association meetings, New York City, 9 April 1954.
16. *Joint Armed Forces statistical classification and basic diagnostic nomenclature of diseases and injuries with a list of surgical operations.* SR-40-1025-1. April, 1949.

RORSCHACH SCORER RELIABILITY[1]

RICHARD H. DANA

St. Louis State Hospital

PROBLEM

Over a thirty year period the Rorschach technique has enjoyed an ever increasing frequency of application within a continually expanding clinical and research framework. Among the most important and least investigated aspects of Rorschach testing is the problem of reliability. There are at least two kinds of reliability germane to projective instruments including (a) reliability of the instrument and (b) reliability of the scorer. The literature contains discussions and research on reliability of the test *per se* but there are few published reports of scorer reliability [1, 6, 7]. In a standard reference, this problem is only briefly mentioned: "It is desirable to make the system of classification as reliable as possible, so that different examiners will arrive at essentially similar classifications of the same sets of responses [8, p. 19]." In a recent overview, Hertz [4] does not mention reliability. Considering the importance of individual scores in the psychogram and the ratios calculated, the problem of scorer reliability should have received more attention in the past.

The paucity of research in this area may be due to an unverbalized awareness that conventional reliability measures may not be applicable to projective materials. A beginning has been made in discussions of reliability in TAT scoring [2], content analysis [8], and more generally, in projective techniques [3]. It is sufficient to indicate here that per cent of agreement, which involves no assumptions, may have more applicability than the usual product moment or tetrachoric correlations between scores of two independent scorers.

METHOD

A representative sample of Rorschach records for each of eleven examiners was obtained by (a) following an alphabetized list of examiners, one record was pulled for each examiner from consecutively numbered files (numbered by admission date) until eleven records had been drawn; (b) this process was repeated six times for a total sample of sixty-six records; (c) this method permitted sampling from all the records in the files for a four year period.

These records were scored independently by two scorers of approximately equal training and experience in other locales, using the Klopfer method [5]. Per cent of agreement was applied as follows: (a) location, determinants, content, and popular main scores were treated as separate items. For example, the response W M H P would consist of four separate items; (b) the total N of separate items present in the six scored records for each examiner was computed; (c) using the N of examiner items as criterion, comparisons were made for presence or absence of identical items in the scoring of the independent scorers.

RESULTS

Thirty-three per cents of agreement were computed (Table 1): Scorer I, Scorer II, and Scorer I with Scorer II, for each of the eleven examiners.

---

[1]This investigation was supported by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service. Principal Investigators are: A. G. Ossorio and A. K. Busch. Research Associate: R. H. Dana. Acknowledgement is made to Jerome Pauker and Jack Werboff, Research Assistants, who did the reliability scoring.

TABLE 1. RORSCHACH PER CENTS OF AGREEMENT BETWEEN TWO SCORERS AND ELEVEN EXAMINERS.

Examiners	N possible agreements	Scorers I	II	I, II
1	235	80	86	86
2	455	73	76	73
3	324	77	84	82
4	414	69	74	75
5	328	71	77	77
6	394	74	74	82
7	291	64	74	75
8	425	71	75	69
9	299	77	78	82
10	409	73	74	75
11	307	69	75	73

DISCUSSION

There was a mean difference of 4.4 points between the agreements of Scorers I and II. This difference is attributable to *degree of care exercised* and use of *personalized scoring variables*. Inspection of all scores indicated that these two kinds of errors, on the part of both the original examiners and the reliability scorers, account for most of the disagreement present.

A figure of 75 per cent of agreement can be considered representative of scoring reliability in this hospital situation. The sample of Rorschach records was scored with adequate reliability to justify inclusion of the entire population of Rorschachs by these same examiners in any research requiring use of Rorschach scoring categories. It is suggested that this is an economical method of determining whether or not there exists a need for rescoring when hospital Rorschach records are used for research purposes.

This method may not have been applied previously because of the implicit assumption that there is only one *right* way of scoring each response and that equally well trained scorers will all score in exactly the same manner. The eleven original examiners in this study had from one to seven years of Rorschach experience; the reliability scorers each had four years of experience. There were no significant relationships between years of experience and reliability of scoring.

The only other use of per cent of agreement for Rorschach reliability reported an overall figure of 76 per cent using an N of 300 responses, five internationally renowned scorers, and a criterion of agreement by three scorers[7]. This striking agreement between experts and the present hospital examiners suggests that increase in training and experience beyond a certain degree of competence does not appreciably affect scorer reliability. If future studies replicate a figure of approximately 75 per cent of agreement for Rorschach scoring, then the 25 per cent of disagreement may be indicative of ambiguity of the scoring categories or of personalized scoring., i.e., influences of examiner personality.

SUMMARY AND CONCLUSIONS

1. Scorer reliability for two independent scorers with each of eleven Rorschach examiners was computed by means of percent of agreement.

2. The thirty-three agreement figures ranged from 64 per cent to 86 per cent with 75 per cent representative of this clinical situation.

3. These per cents of agreement appear sufficiently high for research purposes and the extent to which they fall short of optimal agreement may be a function of ambiguous scoring categories or personalized scoring.

4. Scorer reliability is considered an economical method of determining the need for rescoring Rorschach records to be used in research.

5. The usefulness of per cent of agreement as an index of scorer reliability appears to have been overlooked in objectification of the Rorschach and other projective instruments.

Bibliography

1. Ames, L. B., Learned, J., Metraux, R. W. and Walker, R. N. *Child Rorschach Responses.* New York: Paul B. Hoeber, Inc., 1952.
2. Dana, R. H. Clinical diagnosis and objective TAT scoring. *J. abnorm. soc. Psychol.*, 1955, *50*, 19-25.
3. Dana, R. H. The objectification of projective techniques: Rationale. *Psychol. Rep.*, (In press).
4. Hertz, Marguerite. The Rorschach: thirty years after. In Brower, D. and Abt, L. (Eds.) *Progress in Clinical Psychology.* Vol. I. New York: Grune & Stratton, 1952.
5. Klopfer, B., Ainsworth, Mary, Klopfer, W. and Holt, R. *Developments in the Rorschach Technique.* Vol. I. Yonkers-on-Hudson: World Book Co., 1954.
6. Ramzy, I. and Pickard, P. M. A study in the reliability of scoring the Rorschach ink blot test. *J. Genet. Psychol.*, 1949, *40*, 3-10.
7. Sicha, K. and Sicha, M. A step towards the standardization of the scoring of the Rorschach test. *Ror. Res. Exch.*, 1936-37, *1*, 95-101.
8. Spiegelman M., Terwilliger, C. and Fearing, F. The reliability of agreement in content analysis. *J. soc. Psychol.*, 1953, *37*, 175-187.

THE PERFORMANCE OF HOSPITALIZED PSYCHIATRIC PATIENTS ON THE EGO-STRENGTH SCALE OF THE MMPI

HERBERT QUAY

Milledgeville (Georgia) State Hospital

Problem

The Ego-Strength (Es) scale of the Minnesota Multiphasic Personality Inventory (MMPI) was originally developed by Barron [1] to predict the response of psychoneurotic patients to psychotherapy. However, consideration of the scale suggested a somewhat broader interpretation. Both in his original and later research [1, 2], it was found to be associated positively with aspects of effective personality functioning which he termed "ego-strength". These included intelligence and intellectual functioning, tolerance and a lack of ethnic prejudice, self-confidence, secure sense of reality, independence of judgment, poise, breadth of interest, originality, and personal stability. Negative correlations were found with the clinical scales of the MMPI. This finding was partially attributed to an overlap of items (generally scored in the reverse direction). However, the overlap itself was interpreted as testifying to the general lack of chronic psychopathology among high-scorers.

In a factor-analytic comparison of the Rorschach and the MMPI, Williams and Lawrence [4] found that the *Es* scale has positive correlations with such Rorschach categories as *W*, *M*, *P*, *CF*, and *Fc*. Their factor analysis yielded, among others, a bipolar factor with positive loadings on MMPI *Es* and *K*, Rorschach *W*, *CF*, *Fc*, and *K*, Wechsler-Bellevue Verbal IQ, and negative loadings on MMPI *Pt* and *Sc*. They tentatively labeled this factor as "ego-strength".

With the exception of intelligence, the adjectives of personality description found to be positively associated with high scorers on the *Es* scale appear to be the antipoles of those commonly associated with psychiatrically disabled personalities. The hospitalized psychiatric patient would seldom be characterized as showing a general absence of chronic psychopathology, irrespective of his diagnosis.

The purpose of this study was to investigate the performance of hospitalized psychiatric patients on the *Es* scale as compared to two "normal" or control groups. It was hypothesized that the patients would score significantly lower than either of the two control groups.

Procedure

The patient group (N = 74) was drawn from first admissions to the white women's service of the Milledgeville State Hospital over a five month period. Selection was made on the basis of age (16-50), the presence of sufficient intellectual ability to comprehend the MMPI, the ability to cooperate in a psychometric evaluation, and the absence of demonstrable organic pathology. It should be noted that these selective factors tended to eliminate some of the more highly disturbed patients. However, it was felt that the patients chosen were representative of those new admissions to the service who were capable of participating in a psychometric evaluation which includes the MMPI. The diagnoses of the patients, although they were not separated in the analysis, were distributed in the three broad classifications as follows: 45 psychotics, 22 psychoneurotics and seven conduct disorders.

The patients received the MMPI soon after admission as part of a routine test battery. Thirteen patients who met the aforementioned criteria gave what appeared to be invalid records on the basis of the use of the validity scales as suggested by Meehl [3]. These records were eliminated.

Control subjects were of two groups: 92 affiliate student nurses and 41 psychiatric attendants, all white females. Both groups volunteered to take the MMPI for research purposes, without their names appearing on their papers. Practical limitations did not permit the matching of the groups for age, intelligence, education, and socio-economic status. For such comparison as may be desirable, means and *SD's* for age and education may be found in Table 1. Actual educational levels were not obtained for the student nurses, but all were at least high school graduates.

Table 1. Comparison of the Three Groups for Age and Education

Group	Factors	Age	Education
Patients	Mean	34.19	10.90
	SD	9.18	3.87
	Range	16–49	4–17
Nurses	Mean	20.44	*
	SD	2.82	
	Range	18–34	
Attendants	Mean	42.60	8.76
	SD	5.13	1.80
	Range	24–66	5–13

*Actual educational level not obtained. All 12th grade or better.

Results

Means and S.D.'s for the *Es* scale scores were computed for all groups and may be found in Table 2. To determine the statistical significance of the difference between means, the *t* test was used.

Table 2. Means, Sd's and Ranges of Scores Obtained on the Es Scale for the Three Groups

Factors	Patients	Nurses	Attendants
Mean	37.49	45.06	40.14
Range	21–51	23–54	30–48
SD	7.20	5.77	3.84

Significance of the *t* values found in Table 3 indicates that the hypothesis was substantiated in both parts: the patient group scored significantly lower than either the student nurses or the attendants.

TABLE 3. VALUE OF *t* FOR DIFFERENCES BETWEEN MEANS

Groups compared	*t* Value	Level of Sig.
Patients - Attendants	2.57	5%
Patients - Nurses	7.35	1%

DISCUSSION

From the results, it appears that the scale is measuring factors which are positively associated with functioning, rather than disabled personalities. That it may also be sensitive to differences between "normal" groups is suggested by the differences in means obtained by the two control groups. A *t* computed here gives a value of 5.78 which is significant at the 1% level. Further research is indicated. A comparison of various diagnostic categories would be valuable in view of the varying degree of personality disorganization found in different groups.

SUMMARY

The performance of a group of hospitalized white female psychiatric patients on the Ego-Strength scale of the MMPI was compared to two control groups. The 74 patients were found to score significantly lower than 92 affiliate student nurses and 41 psychiatric attendants.

REFERENCES

1. BARRON, F. An ego-strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, 17, 327-333.
2. BARRON, F. Ego-strength in normal men as measured by an MMPI scale. Paper read at the Western Psychol. Ass., 1954.
3. MEEHL, P. E. Profile analysis of the Minnesota Multiphasic Personality Inventory in differential diagnosis. *J. Appl. Psychol.*, 1946, 30, 517-524.
4. WILLIAMS, H. L. and LAWRENCE, J. F. Comparison of the Rorschach and MMPI by means of factor analysis. *J. Consult. Psychol.*, 1954, 18, 193-197.

USE OF EYE MOVEMENTS AS A RESPONSE INDICATOR IN TESTING THE PHYSICALLY HANDICAPPED*

DONALD E. WAMBA AND STANLEY S. MARZOLF
Public Schools, Corona, California *Illinois State Normal University*

PROBLEM AND METHOD

In order to evaluate abilities of the handicapped one must choose tests which require a minimal use of response mechanisms or the use of that mode of response which is least disabled. In many instances neither verbal nor performance tests can be used because speech is unintelligible and coordinated hand and arm movements are virtually impossible. In such cases it is seldom that the use of extra-ocular muscles is so impaired as to make coordinated eye movements unrecognizable. Even when nystagmus exists it is usually possible to discern the locus of regard.

To make use of eye fixation as a means of indicating choice of response to test situations, a box was constructed which permits the examiner to observe the subject's eye movements when instructed to look at his choice of correct test-item response. A box 9 x 9 x 12 inches was made of sheet metal. In the lower part of one of

*Adapted from a thesis presented by Mr. Wamba in partial fulfillment of the requirements for the degree of Master of Science in Education, Illinois State Normal University.

the square sides of this box a 6 x 8 inch opening was cut. About this opening a frame was so constructed that a sheet of transparent material could easily be inserted and removed. Below the opening a bracket was attached for holding test material. A viewing aperture, made from the viewing chamber of a stereoscope, was placed in the opposite side of the box. The entire apparatus was mounted on a tripod.

To use the apparatus the "problem" portion of a test item is placed on a card and inserted in the rack below the open side of the box. The alternative possible responses, or foils, for that item are pasted in a circular arrangement, equidistant from one another, on a transparent sheet which is then inserted in the frame before the large opening in the box. The box is placed before the subject's eyes at a distance of approximately 10 inches. The examiner, with his back to the light source of the room, gives the instructions and looks through the viewing aperture to note which of the foils the subject fixates upon when told to "look at" the correct choice for the problem presented.

A brief study was carried out to determine whether it is reasonable to suppose that a test administered by the procedure just described will yield results which closely approximate those obtained by the standard mode of test administration. The test chosen for this purpose was the *Progressive Matrices*.[1] Tracht[(2)] has considered this test useful for testing the severely handicapped even when administered in the standard manner.

The "problem" portion of each item of the *Progressive Matrices* was cut from the booklet and pasted on a card. The answer foils were likewise cut from the booklet and pasted on a sheet of plastic in a circular arrangement equidistant from one another. For each item there is a separate card and a separate plastic sheet. The procedure for administering the test described by Raven was followed except that whenever the directions call for "pointing to" the subject was asked to "look at."

Forty-six children ranging in age from six to 11 years, without physical handicap, were randomly assigned to one of two experimental groups, A and B. Group A was first given the test in the usual booklet form and approximately one week later the above-described adaptation. For Group B the order of giving the two forms was reversed.

Results and Conclusions

The mean scores for the two forms are identical, 22.46, but the standard deviation for the booklet form is 3.32 while the standard deviation for the adaptation is only 2.62. However this difference in variability is not significant at the five-per-cent level. The correlation between the two forms for Group A is .86 $\pm$ .05 and for Group B., .85 $\pm$ .06. The correlation between first testing and second testing, Groups A and B combined, is .88 $\pm$ .03.

Table 1. Analysis of the Variance in Scores

Source	df	Sum of squares	Variance
Between forms	1	0	
Between orders	1	636.566	636.566
Interaction	1	136.347	136.347
Within sets	88	2,569.914	29.204
Total	91	3,342.827	

Application of the F-test to the results in Table 1 indicates that the difference due to order of presentation is significant at the one-per-cent level and that the interaction variance is significant at better than the five-per-cent level. (There is a small but fairly reliable advantage in practice effect when the adaptation form is given first.) Most important of course is the lack of any difference between the two forms of the test.

Though these data are limited in scope, if we may assume randomness of sample, it appears that tests may be given using eye movements as indicators of response without altering the quantitative findings to a significant degree. Use of the procedure is not limited to the *Progressive Matrices* but may be extended to other kinds of tests, verbal or non-verbal, so long as responses may be indicated by a choice reaction.

References

1. Raven, J. C. *Guide to Using Progressive Matrices (1947): Sets A, Ab, B.* London: H. K. Lewis, 1949.
2. Tracht, Vernon S. "Preliminary findings on testing the cerebral palsied with Raven's Progressive Matrices." *J. except. Children*, 1948, 15, 77-79.

THE EFFECT OF PREFRONTAL LOBOTOMY UPON ABSTRACT BEHAVIOR*

ALBERT R. MASLOW

Catholic University of America

Problem

Does psychosurgery result in an intellectual loss? Klebanoff, Singer and Wilensky [8] have written an excellent summary of the literature related to this question. Despite considerable inconsistency, the studies reviewed are suggestive of an impairment in functioning which is described either as intellectual or, more specifically, in abstract behavior. Much of the conflict in results is no doubt due to faulty experimental procedure, particularly by the absence of adequate control groups.

The purpose of this study was to test, in a more precise manner than previously attempted, Goldstein's position [3] that prefrontal lobotomy causes an impairment in abstract behavior. A special adaptation of the Holsopple Concept Formation Test which involves the concepts of "shape," "color," "numeral," and "number of figures" was used in this study. In accordance with Goldstein's views, it was hypothesized that the lobotomized patient would show an impairment in his ability to form concepts, and that the degree of impairment would be proportionate to the degree of abstractness of the concept. Tying this in with Heidbreder's results (when working with drawings) concerning levels of concreteness [4, 5, 6], it was predicted that prefrontal lobotomy would result in a decrease in the ability to form the concepts of "numeral" and "number of figures", and little or no decrease in the ability to form the concept of "shape." Because of Heidbreder's negative findings in working with color, no predictions were made regarding the relative difficulty of this concept following the operation.

Method

The subjects used in the present study were 40 chronic schizophrenics taken from six VA hospitals. The 16 who received the standard Freeman-Watts prefrontal lobotomy constituted the experimental group, and the remaining 24 served as controls. All subjects were given Holsopple's test pre- and three months postlobotomy. In order to compare the two groups, they were matched separately into 16 pairs for the total test and for each of the four concepts included in the test. In each case the

---

*This paper is adapted from a thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology and Psychiatry at The Catholic University of America. The author is indebted to Dr. John W. Stafford and Dr. Maurice Lorr for their guidance and support, and also for the contributions of Dr. J. Q. Holsopple, Principal Investigator of the VA Lobotomy Research Project, for his help in this study. The apparatus was the product of many years of Dr. Holsopple's research. The raw data were made available without identification of the patients involved, through Dr. Holsopple's efforts, and his counsel was available from time to time.

matching was based primarily upon initial performance and secondarily upon overall preoperative severity of illness score. The severity of illness score was derived from ratings by psychiatrists on the Multidimensional Scale for Rating Psychiatric Patients, devised by Lorr and Jenkins[7, 9]. In addition, the 24 controls and the 16 operated were matched as groups on the basis of age and duration of illness (i.e., length of time in last hospitalization). Following these procedures, the tests were compared to determine whether there was a differential change in performance between the groups from pre- to three months postlobotomy for the total test and for each concept.

Results

In comparing scores between the experimental and control groups, the method used was that suggested by McNemar[10, pp. 78-82] in which two groups are matched and the significance of a change in one group following an interpolated experience is to be determined. From Table 1, there is no significant difference between the groups

Table 1. Significance of the Differences (Dd) Between the Experimental and Control Groups Regarding Change in Mastering the Concepts

Factors	Dd	σDd	C.R.	P
Shape	.187	.207	.90	> .05
Color	−.126	.198	.64	> .05
Numeral	−.126	.153	.82	> .05
No. Figures	.063	.165	.38	> .05
Total Test	.062	.108	.57	> .05

on any of the four concepts or on the total test. Thus, a change in ability to master the concepts on the test as a result of the operation is not indicated.

In further analysis of the data, it was considered desirable to compare the subjects on the basis of the concepts influencing their sorting, independently of their mastery of the problem concepts. Following Cronbach's suggestion[1, 2], a method of dividing the population into sub-groups and matching for size of denominators was used. The denominator represented the total number of concepts involved in the initial responses in each problem (excluding the problem concept). Of the 12 critical ratios calculated, only one had a P value significant below the .05 level of confidence. This was a C.R. of 2.82 which was significant at the .01 level for a difference in the two groups in response to the concept of number of figures when the problem was numeral.

In order to better evaluate this value, the significance of change in percentage was calculated separately for the two groups. For the operated and control groups, the C.R.'s were 1.46 and 1.44 respectively. Neither value is statistically significant. Hence, since our other evidence showed no indication of meaningful changes in the control group, it seems unreasonable to attach meaning to the above finding. Also, in view of the number of statistical computations that were made plus the numerous possible calculations that were implicitly rejected by observation, we might expect at least one apparently significant finding to occur by chance alone.

Summary and Conclusions

The present study was undertaken to test the hypothesis that prefrontal lobotomy causes an impairment in the ability to form concepts, and that the degree of impairment is proportionate to the level of abstractness of the concept. The subjects consisted of 40 chronic schizophrenics, of whom 16 received the Standard Freeman-Watts Prefrontal Lobotomy. The remaining 24 were used as controls. All subjects were given a modified version of the Holsopple Concept Formation Test before and three months after the operation period. Statistical analyses were made to determine whether there were significant differential changes in the performance of the operated group. No significant differences between the groups were found.

Our conclusion, therefore, is that we have no evidence to indicate that the prefrontal lobotomy affects abstract behavior, and further that the standard lobotomy operation, when administered to chronic schizophrenics, does not cause a general impairment in ability to form concepts.

References

1. Cronbach, L. J. The reliability of ratio scores. *Educ. Psychol. Measurement*, 1941, 1, 269-278.
2. Cronbach, L. J. Statistical methods applied to Rorschach scores: A review. *Psychol. Bull.*, 1949, 46, 393-429.
3. Goldstein, K. Frontal lobotomy and impairment of abstract attitude. *J. Nerv. Ment. Dis.*, 110, 1949, 93-111.
4. Heidbreder, E. The attainment of concepts: I. Terminology and methodology. *J. Psychol.*, 1946, 35, 173-189.
5. Heidbreder, E. The attainment of concepts: II. The problem. *J. Psychol.*, 1946, 35, 191-223.
6. Heidbreder, E. The attainment of concepts: III. The process, *J. Psychol.*, 1947, 24, 93-138.
7. Jenkins, A. L. and Holsopple, J. Q. Criteria and experimental design for evaluating results of lobotomy. Chapter 27 of *Psychiatric Treatment*. Proced. Assoc. Nerv. Ment. Dis., 1953.
8. Klebanoff, S. G., Singer, J. L. and Wilensky, H. Psychological consequences of brain lesions and ablations. *Psychol. Bull.*, 1954, 51, 1-41.
9. Lorr, M. Multidimensional Scale for Rating Psychiatric Patients. *VA Techn. Bull.* 10-507, (Nov. 1953), 1-44.
10. McNemar, Q. *Psychological statistics*. New York: Wiley, 1949.

A FOLLOW-UP ITEM ANALYSIS OF THE TAYLOR ANXIETY SCALE

ARNOLD H. BUSS

Larue Carter Memorial Hospital, Indianapolis

Problem

Hoyt and Magoon [1] recently reported a validation study of the Taylor Anxiety Scale. Using counselors' judgments of manifest anxiety in college students as the criterion, they performed an item analysis on the Taylor Anxiety Scale. It was found that 30 items significantly differentiated High Anxious from Low Anxious clients. The present study is an attempt to extend the item analysis to a patient population.

Procedure

Neuropsychiatric patients were interviewed by a clinical psychologist, with three other psychologists observing.[1] The four psychologists rated the degree of manifest anxiety on a five-point scale. "Anxiety" included most of the behaviors used by Hoyt and Magoon in their study: restlessness, motor tension, embarrassment, and physiological concomitants (sweating, flushing, etc.). Sixty-four patients were interviewed and subsequently administered the Taylor Anxiety Scale. The patients varied in age, sex, and diagnosis, and they represent a diverse sample of the psychiatric in-patient population.

Results and Conclusions

The interview ratings of the four psychologists were pooled. The correlation (Pearson Product-Moment) between these pooled ratings and the Taylor Anxiety Scale scores was .60. This value, taken together with Hoyt and Magoon's contingency coefficient of .57 for college students, suggests that the Taylor Scale has fair validity for diverse populations.

---

[1]The author wishes to thank Dr. Morton Wiener, Mrs. Ann Durkee and Mr. Marc Baer, who acted as observers.

Next, the patients were divided into High Anxious(N=22), Medium Anxious (*N*=22) and Low Anxious (*N*=20) groups on the basis of the pooled interview ratings. An item analysis was performed on the individual Taylor Scale items, using only the High Anxious and Low Anxious groups.

Hoyt and Magoon found 16 items significant at the .05 level of confidence in both of their samples. Of these the following items had Chi-squares that were significant at the .05 level in the patient population.

1. I work under a great deal of tension.[2]
2. I cannot keep my mind on one thing.
3. I frequently find myself worrying about something.
4. I am usually calm and not easily upset.
5. I have periods of such great restlessness that I cannot sit long in a chair.
6. I find it hard to keep my mind on a task or job.
7. I am not unusually self-conscious.
8. Life is a strain for me much of the time.
9. I am certainly lacking in self-confidence.

Hoyt and Magoon found that 14 items significantly differentiated clients in one sample but not in the other sample, although the direction of the difference was the same. Of these 14 items only five had significant Chi-squares in the present study:

10. I have a great deal of stomach trouble.
11. I certainly feel useless at times.
12. I am a high-strung person.
13. Sometimes I feel that I am almost about to go to pieces.
14. I shrink from facing a crisis or difficulty.

Of the items Hoyt and Magoon found not significant in either of their samples, the only one significant in the present study was "I have diarrhea once a month or more." However, the direction of the difference between High and Low Anxious answers on this item was opposite to that found in the previous study.

Thus, in attempting to extend the validity of the Taylor Scale to a patient population, there was considerable "shrinkage." Of the 30 items valid for a college student population, only 14 items are valid for a patient population.

Reference

1. Hoyt, D. P. and Magoon, T. M. A validation study of the Taylor Manifest Anxiety Scale *J. clin. Psychol.*, 1954, *10*, 357-361.

[2]The 1953 revision of the Taylor Scale was used in the present study, and some of the items had been paraphrased. The items presented here have the same wording as those used by Hoyt and Magoon.

THE RELIABILITY OF WITTENBORN'S SCALES FOR RATING CURRENTLY DISCERNIBLE PSYCHOPATHOLOGY[1]

SHIRLEY S. PUMROY AND WILLIAM S. KOGAN

University of Washington — *Veterans Administration Hospital, Seattle, Washington*

Problem

The rating scales recently developed by Wittenborn et al[3, 4] offer a means of describing current symptom patterns in an objective, quantitative form. The rating scales were developed in two-step procedure: first, appropriate behavior items were assembled and screened to meet specified criteria; next, judges using the items rated a large number of hospitalized psychiatric patients. The ratings were correlated and the resulting matrix factor analyzed. A number of symptom clusters were obtained which seemed to follow the characteristics of several of the traditional diagnostic categories. A scoring system was worked out in terms of these clusters.

In the construction of the scales Wittenborn stressed the importance of the scales being relatively independent of the insights and sophistication of the rater, his bias, or theoretical persuasion. Another important factor is the reliability of the scale, which is related to the matter of rater differences. Application of the scale clinically demands "adequate" reliability, and such data have not been available to date. Wittenborn states[3] that the best type of reliability study would be one wherein two different psychiatrists rate the same patient independently for some specified period of observation. This study concerns itself with investigating this method of determining reliability.

Procedure

Eleven closed ward psychiatric patients in a Veterans Administration General Medicine and Surgery Hospital were each rated by four observers: two doctors and two nurses. One doctor was the doctor to whom the patient was assigned for treatment and therefore varied; the other was the doctor responsible for ward management, who rated all patients. One nurse rated all but one patient; the other nurse varied. Seven different nurses were involved in the ratings. The ratings were made at the end of three days of observation by the four raters. A not too successful attempt was made to rate the patients within ten days to two weeks after admission. No attempt was made to select patients with respect to psychiatric diagnosis.

Results and Discussion

Amount of agreement among raters—Symptom clusters. Table 1 indicates that the coefficients of concordance W[2] on five of the symptom clusters were significant above

Table 1. Rater Agreement on Nine Symptom Clusters

Scales	Coefficient of concordance (All raters)	Generalized reliability coefficient (All raters)
Acute anxiety	.25	—†
Conversion hysteria	.41	—
Manic state	.71**	.87
Depressed state	.50*	.66
Schizophrenic excitement	.69**	.85
Paranoid condition	.58**	.76
Paranoid schizophrenic	.71**	.86
Hebephrenic schizophrenic	.63**	.81
Phobic compulsive	.18	—

*chi square significant at the 5% level.
**chi square significant at the 1% level.
†omission indicates failure to reach 5% level.

[1]From Veterans Administration Hospital, Seattle, Washington.
[2]Adapted by Allen Edwards from a technique developed by Horst[1].

the one per cent level. For these five, the generalized reliability coefficient[2] gives an estimate of the reliability with a replication of judges. Agreement on the nine clusters was calculated for the four raters and appears in Table 1.

Further analyses, the data for which are not presented here, revealed that agreement between doctors and between nurses is better than between all raters, that the patient's own doctor tended to rate all scales higher, and that different raters stressed different symptom clusters. For example, in the acute anxiety cluster the doctors felt that anxiety was an important element in these patients while the nurses did not. The nurses emphasized the depressed state cluster while the doctors did not.

Amount of agreement among raters—Individual patients. The agreement among raters on individual patients was also investigated. On only three of the eleven cases did the rater agreement reach the five per cent level of significance.

These data seem to raise some question as to whether Wittenborn's objective of developing a procedure equally applicable to raters of differing levels of psychiatric sophistication is fully realized. In addition, it appears that, for this limited sample, some symptom clusters are more reliable than others. The acute anxiety and phobic compulsive clusters were particularly poor. Detailed interviews with the raters stressed the difficulty in rating items that require the rater to make inferences concerning the internal state of the patient. This is supported by their lack of agreement on such items. It is suggested that further investigations of these probable sources of unreliability will make this approach to quantification more clinically useful. A study of the conditions under which the observers can agree must be made before Wittenborn's type of scale can be used for psychometric purposes.

Summary

Eleven closed ward psychiatric patients were rated by each of four raters on Wittenborn's scales for rating currently discernible symptom patterns. Two doctors and two nurses rated each patient. The results of this reliability study were as follows:

1. Patient's own doctor rated all scales higher.
2. Different raters stressed different symptom clusters.
3. Amount of agreement on individual patients was relatively low.
4. Scales involving inferences as to internal state of the patient showed low reliability.
5. Acute anxiety and phobic compulsive scales were least reliable.

Question was raised as to whether the technique was suitable for use by individuals at all levels of psychiatric sophistication. Possible sources of unreliability were discussed.

References

1. Horst, P. A generalized expression for the reliability of measures. *Psychometrika*, 1949, 14, 21-23.
2. Kendall, M. G. *Rank correlation methods.* London: Charles Griffin & Co., 1948.
3. Wittenborn, J. R. Symptom patterns in a group of mental hospital patients. *J. consult. Psychol.*, 1951, 15, 290-302.
4. Wittenborn, J. R. and Holzberg, J. D. The generality of psychiatric syndromes. *J. consult. Psychol.*, 1951, 15, 372-380.

A RELIABILITY STUDY OF A SHORT FORM OF THE WISC VOCABULARY SUBTEST

RENATE GERBOTH ARMSTRONG

East Moline State Hospital, Illinois

Introduction

With the continually increasing case loads in most hospitals and clinics, time available for testing becomes more and more at a premium. Therefore, "short forms" of the Wechsler Scales, adult and children, have been the subject for many studies. Although some of the studies of the Wechsler Bellevue short forms have shown high reliability,[2, 3, 4] the same has not as yet been found for Short Forms of the WISC, especially when used on a heterogenous group. However, since it has been proved "that relatively little loss in reliability is involved in the use of either of the alternate word lists as a substitute for the full vocabulary list"[1] on the Wechsler Bellevue, it is the purpose of this study to determine whether, by the use of the split half method, the vocabulary subtest of the WISC would still prove as reliable as the total list for an intellectually heterogeneous population of children referred for numerous reasons to a child guidance clinic.

Procedure

Since the psychologists from the East Moline State Hospital are participating in the testing program of the children referred to the Rock Island County Child Guidance Clinic, test data from this clinic were used for this study. Two hundred full scale WISC records from the files were selected so as to construe an even distribution as to sex and age. One hundred male and one hundred female children were selected from five age ranges of two years each, so that the overall age range covered extended from 5-0 to 14-11 years. The scoring of all vocabulary responses was checked by the author. Then the vocabulary list was split into halves, the odd words composing one and the even words the other half. Raw scores on these two lists were computed for each record and recorded in one of the ten age and sex groups, composed of twenty records in each. Then the raw scores on the two lists were correlated, using the Pearson product-moment correlation method, first for each of the ten groups and secondly for the total population and probable errors for the correlations were also computed. In addition to means and standard deviations of each group, the Spearman-Brown formula for estimating reliability from two comparable halves of the test was utilized.

Results

Means of raw scores for the different age groups of odd items range between 8.60 to 20.50 and even items between 6.85 to 19.20. This wide range is, of course, due to the increase in verbal abilities with age. The overall mean for odd items was 14.62 and for even items 13.32. The standard deviations of 5.29 and 5.62 are somewhat higher than would be expected in a more normal group of subjects, but in this study can be attributed to the heterogeneity of a clinic population. Correlations between odd and even items fall between .76 and .93, the higher correlations occurring mostly in the lower and higher age groups. The overall correlation of .88 is, as are the others, significant at the 1% level of confidence. None of the probable errors exceed 0.07, indicating that it is quite unlikely that the true r's vary by more than $\pm.07$. The Spearman-Brown formula reveals that the estimated reliability of the split half vocabulary for the ten groups ranges from .87 to .96, the total sample yielding a reliability of .94. Therefore, it is apparent that relatively little loss in reliability occurs in the use of either of the alternate word lists as compared to the use of the complete list.

TABLE 1. MEANS AND STANDARD DEVIATIONS FOR EACH HALF OF THE VOCABULARY SUBTEST, CORRELATIONS OF THE SUBTEST HALVES, PROBABLE ERRORS OF THE CORRELATIONS, AND ESTIMATED RELIABILITY COEFFICIENTS

Age Group	Odd Items Means	Odd Items S. D.	Even Items Means	Even Items S.D.	Correlations Odd-Even Items	Probable Errors	Reliabilities
Girls							
5–7	8.60	3.44	6.85	3.38	.81	0.05	.90
7–9	11.70	2.28	10.20	4.17	.83	0.04	.91
9–11	14.10	3.05	12.45	4.28	.79	0.06	.89
11–13	15.90	3.67	14.75	4.28	.78	0.05	.88
13–15	15.95	5.28	14.80	5.11	.93	0.06	.96
Boys							
5–7	11.45	3.39	10.50	2.45	.84	0.04	.92
7–9	11.90	5.07	10.95	4.45	.86	0.04	.93
9–11	16.95	3.81	15.05	3.10	.76	0.07	.87
11–13	19.20	3.72	18.85	5.13	.77	0.06	.88
13–15	20.50	3.52	19.20	3.76	.81	0.05	.90
Total	14.62	5.29	13.32	5.62	.88	0.03	.94

CONCLUSION

A split half reliability study was conducted in order to determine the feasibility of using a short form of the WISC vocabulary subtest for a child guidance clinic population. The vocabulary raw scores from two hundred records of children referred to the Rock Island County Child Guidance Clinic were selected with even distribution according to age and sex. Correlations ranged from .76 to .93 with an overall correlation of .88 for the total group. None of the probable errors exceeded 0.07, and all correlations were significant at the 1% level of confidence. Also, the Spearman-Brown correction formula yielded a reliability coefficient of .94 for the complete sample and a range of .87 to .96 for the ten groups. These results suggest that the loss of reliability involved in using either of the alternate word lists, instead of the total vocabulary list, is minimal, especially when compared to the time saved. Nevertheless, further research is needed to establish the complete reliability of the short form vocabulary subtest in order to use it for more heterogenous types of children.

BIBLIOGRAPHY

1. FINKELSTEIN, M., R. GERBOTH, and R. WESTERHOLD. Standardization of a Short Form of the Wechsler Vocabulary Subtest. *J. Clin. Psychol.*, 1952, 8, 133-135.
2. HUNT, W. A. et al. The validity of some abbreviated intelligence scales. *J. consult. Psychol.*, 1948, 12, 48-52.
3. MCNEMAR, QUINN. On abbreviated Wechsler-Bellevue Scales. *J. consult. Psychol.* 1950, 14, 70-80.
4. PATTERSON, C. H. A comparison of various "short forms" of the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 1946, 10, 260-67.
5. WECHSLER, D. *The measurement of adult intelligence (3rd Ed.)* Baltimore: Williams and Wilkins, 1944.
6. WECHSLER, D. *Wechsler Intelligence Scale for Children: Manual.* New York: The Psychological Corporation, 1949.

NOTE ON HYPOTHETICAL SITUATIONS IN PERSONALITY APPRAISAL

HERMAN FEIFEL

Veterans Administration Mental Hygiene Clinic, Los Angeles, California

The potentiality of hypothetical situations for permitting expression of needs, goals, and values has been emphasized by Lehner and Saper[2]. Since the situations allow a variety of responses, albeit "anchored to some extent in the awareness of what is expected," it is felt they afford a rapid method of probing certain covert aspects of personality. The present note deals with the responses of a psychoneurotic and psychotic population to two hypothetical situations and compares them with reported group data on "normals."

Method

The subjects consisted of two groups of mentally ill patients in a Veterans hospital. One group comprised 40 acutely disturbed closed ward patients in partial remission; the other, 51 open ward patients diagnosed as psychoneurotic. Paranoid schizophrenics (80%) dominated the closed ward population, and anxiety (36%), depressive (28%, and obsessive compulsive (22%) reactions characterized the open ward patients. Both groups were well matched on the factors of age, intelligence, education, and occupational background. Mean ages were 33 and 35 years; average IQ's were 105 and 108; educational level was about three years of high school; and job backgrounds consisted mostly of clerical and skilled trades.

Both groups of patients were asked to answer the two following hypothetical situations: (a) If you were told you could have any three wishes come true, what would you wish? (b) What do you believe to be three deeds which you consider yourself incapable of doing? Scoring followed the procedure suggested by Lehner and Saper.[2]

Results and Discussion

Three Wishes. Both groups of patients selected health (32%; 30%) as their primary wish. This is certainly not unexpected for hospitalized patients. Happiness, wealth, and marital fulfillment, respectively, came next. Significant differences at the .05 level were found between the groups with regard to social welfare and achievement which were the two least chosen categories. The closed ward patients chose the social welfare category reliably more often than the open ward patients probably reflecting their counter-reaction to greater aggressive promptings and self-held passive wants. The open ward patients, on their side, selected the achievement category significantly more often than the closed ward patients. This was most likely related to their closer proximity to the outside world with its demands. Reliability of the categories was determined by having an independent judge score the responses. Agreement in classification was 91 per cent.

Although the groups are not quite comparable, comparison of the wishes of our patients with the college women of Wilson[3] and student population of Lehner and Saper[2] showed a good deal of similarity. Self-interest wishes generally predominated over those of an altruistic nature. The trend in the wish categories supported the idea that persons usually wish for things which experience has taught them are both valuable and attainable[3]. Also, to a large degree, the answers manifested feelings and motives which could not be considered "solely personal" and representative of the "deeper layers" of personality functioning.

Three Impossibilities. Both groups of patients selected the achievement category as their main area of incapability, 48 per cent of the closed ward and 66 per cent of the open ward patients doing so. This was followed by hurting others, manipulating one's surroundings (e.g. "I could never change my wife;" "I don't think I could stop

war," etc.), and murder in the closed ward group, and by manipulating one's surroundings and dishonesty in the open ward patients. The intense preoccupation with achievement, in both groups, highlights the cultural influence on men in this field. The open ward patients chose the achievement category significantly more often at the .01 level than the closed ward patients, reflecting both their relative proximity to life demands outside the hospital and doubts concerning their coping abilities. The grandiosity and irreality of some of the more acutely disturbed patients was evident in the answers of four of them, "there's nothing I can't do."

Comparison of our results with those of Lehner and Saper indicated little difference. The students' major category choices were respectively achievement, manipulation of surroundings, murder, and dishonesty. Reliable differences, however, have been cited relating to sex and level of intelligence[1, 2]. Analysis of individual responses added weight to the contention that "the impossibility, whenever it is subjectively colored, represents an overestimate of either a wish or a fear"[1].

Summary

1. Health and happiness dominated the wish choices of our patients. They were more concerned with their own personal problems than those of the world. By and large, they wished for things that experience had taught them were within attainable reach.

2. With regard to impossibilities, the achievement motive prevailed in both groups, significantly more so for the open ward patients. This probably mirrored their relative nearness to responsibilities outside the hospital and attendant concern over handling them.

3. There was little difference between the wish and impossibility choices of our patients and those of "normal" groups. The modal tendencies revealed by the patients suggest, at least based on the two situations presented in this paper, that the hypothetical situation technique reflects cultural and conventional stereotypes more than it does personal and idiosyncratic outlook. It does, however, permit the establishing of group reference points from which baseline the clinician can assess individual responses and deviations. Increased fertile use of the technique will have to provide expression for more specific individual meanings related to the person's coping and defensive responsiveness to life situations.

References

1. Diamond, S. Three impossibilities: a verbal projective technique. *J. Psychol.*, 1947, *24*, 283-292.
2. Lehner, G. F. J., & Saper, B. Use of hypothetical situations in personality assessment. *J. Pers.*, 1952, *21*, 91-102.
3. Wilson, F. T. Verbally expressed wishes of children and college women students. *J. Psychol.*, 1938, *5*, 91-105.

THE DEVELOPMENT OF SEXUAL DIFFERENTIATION ON THE DRAW-A-PERSON TEST[1]

CLIFFORD H. SWENSEN AND KENNETH R. NEWTON

University of Tennessee

In a previous paper [3] the senior author described a scale developed for rating sexual differentiation on the Draw-a-Person Test. It was felt that it would be desirable to trace the development of sexual differentiation in children, as it is revealed in their human figure drawings.

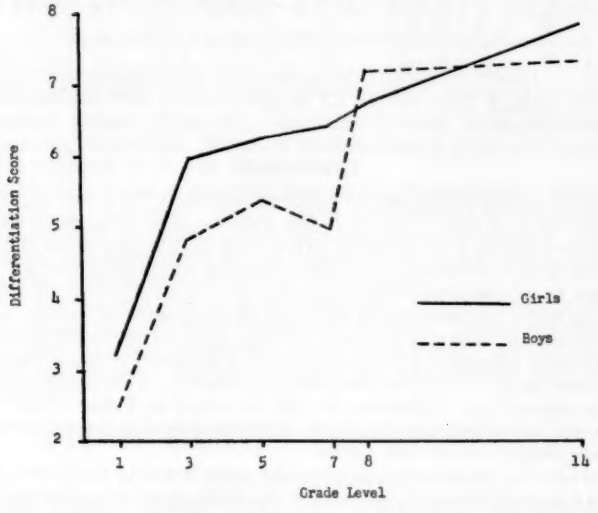
Procedure

The Draw-a-Person Test was administered as a group test to 163 children in the first, third, fifth, seventh, and eighth grades, and to twenty-two college students in an elementary psychology class. The sex of the first-drawn figure was noted, and the pairs of drawings for each person were rated for sexual differentiation using the scale described in a previous paper [3].

Results

As can be seen in Figure 1, sexual differentiation improves in a negatively accelerated curve, as a function of age. And, as is shown in Table 1, girls tend to differentiate between the sexes significantly better than boys. However, at about age thirteen the boys' sexual differentiation catches up with that of the girls, and beyond age thirteen there is apparently no significant difference between males and females in sexual differentiation on the Draw-a-Person Test.

Figure 1. Sexual Differentiation on the Draw-a-Person Test as a Function of School Grade



Some rather unexpected findings occurred with respect to the sex drawn first. As can be seen in Table 1, the first grade girls drew their sex first in 91 per cent of the cases. Beyond the first grade the percentage of girls drawing their own sex first dropped until it reached a low of 72 per cent for the fifth grade girls. Beyond the fifth grade the girls tended to draw their own sex in about 75 per cent of the cases. The boys showed a somewhat different pattern. The first grade boys drew their own

[1]The authors wish to express their appreciation to Sister Mary Noreen, C. S. M., Principal of St. Mary's School, Knoxville, for her assistance in obtaining drawings.

sex first in 86 per cent of the cases, and this percentage increased to 100 per cent for the seventh grade boys. In the eighth grade only 73 per cent of the boys drew their own sex first, but the male college students drew their own sex first in 89 per cent of the cases.

TABLE 1. SEXUAL DIFFERENTIATION RATINGS BY SCHOOL GRADE AND SEX

School Grade	Sex	N	Mean Age	% drawing own sex first	Mean sexual differentiation rating	t of differences between sexes on sexual differentiation
1	Male	21	6.14	86	2.57	2.69*
	Female	22	6.32	91	3.27	
3	M	21	8.22	86	4.86	3.21**
	F	16	8.29	88	5.94	
5	M	15	10.40	93	5.40	2.12*
	F	14	10.07	72	6.29	
7	M	7	12.00	100	5.00	2.58*
	F	17	12.29	76	6.47	
8	M	11	13.27	73	7.18	1.15
	F	19	13.00	79	6.79	
college sophomores	M	9	20.56	89	7.33	.95
	F	13	19.00	77	7.85	

*significant at .05 level
**significant at .01 level

DISCUSSION

These results indicate that sexual differentiation on the Draw-a-Person Test increases as a function of age. However, the pattern of development seems to differ between males and females. Girls seem to improve in differentiation chiefly from age six to eight, and girls differentiate between the sexes significantly better than boys in the years prior to adolescence. Boys appear to improve in differentiation rapidly from ages six to eight, progress more slowly from ages eight to twelve, then improve quite rapidly from age twelve to thirteen. At age thirteen the boys are differentiating between the sexes as adequately as the girls are. This spurt of improvement from age twelve to thirteen might be due to the onset of puberty. This explanation would appear to be the logical one. However, it will be noted in Table 1 that we had only seven boys in our seventh grade group, so it is possible that these results are due to an error in sampling.

The trends in the percentages of boys and girls drawing their own sex first were unexpected and somewhat at variance with the findings of other investigators[1, 4, 5]. Weider and Noller[4, 5] in their two studies of boys and girls from ages eight to twelve found that boys tended to draw their own sex first about 70 per cent of the time and girls tended to draw their own sex about 90 per cent of the time. But the most interesting aspect of the percentage drawing their own sex first becomes apparent in comparing the trends in percentage drawing their own sex first with the ratings on sexual differentiation. While there is a fairly regular increase in sexual differentiation with increased age, there is no comparable increase in the percentage drawing their own sex first. Thus, it would appear that increased ability to differentiate between the sexes is not related to a tendency for subjects to draw their own sex first.

SUMMARY

The Draw-a-Person Test was administered to 163 grade school students from grades one through eight and to twenty-two college students. The drawings were rated for sexual differentiation and the sex of the first drawn figure was noted. It was found that sexual differentiation increased with increased age. Up to the eighth grade the girls differentiated significantly better than boys between the sexes. Beyond the eighth grade there was not a significant difference between the sexes in sexual differentiation. Tendency of subjects to draw their own sex first did not appear to be related to increased sexual differentiation on the drawings.

REFERENCES

1. JOLLES, I. A study of the validity of some hypotheses for the qualitative interpretation of the H-T-P for children of elementary school age: I. Sexual identification. *J. clin. Psychol.*, 1952, *8*, 113-118.
2. MACHOVER, KAREN. Human figure drawings of children. *J. proj. Tech.*, 1953, *17*, 85-91.
3. SWENSEN, C. H. Sexual differentiation on the Draw-a-Person Test. *J. clin. Psychol.*, 1955, *11*, 37-41.
4. WEIDER, A. and NOLLER, P. A. Objective studies of children's drawings of human figures. I. Sex awareness and socioeconomic level. *J. clin. Psychol.*, 1950, *6*, 319-325.
5. WEIDER, A. and NOLLER, P. A. Objective studies of children's drawings of human figures. II. Sex age, and intelligence. *J. clin. Psychol.*, 1953, *9*, 20-23.

THE HOWARD INK BLOT TEST

A NEW PROJECTIVE TECHNIQUE

- **Derived experimentally by James W. Howard, Ph.D.**
- **An individual test consisting of 12 plates 8½" x 11". The blots are somewhat larger than the Rorschach and with a wider range of stimulating features. Blots of varied colors selected from an experimental sample of 10,000 blots.**
- **Not intended to be a parallel test to the Rorschach. Statistical findings from a "normal" group show the test to differ markedly from the Rorschach in ways which should indicate greater sensitivity.**
- **Selection and arrangement of blots made on the basis of evidence of character, degree and spread of stimulation.**
- **Data from normative sample indicate more frequent light determined responses, higher incidence of color responses with different C, CF and FC relationships, more movement responses and fewer animal responses. Greater diagnostic sensitivity with wider range of response.**

Price: $12.50 (In U. S. and Canada; foreign postage $.50 extra)

HOWARD INK BLOT TEST MANUAL

MONOGRAPH SUPPLEMENT NO. 10 FROM JULY 1953 ISSUE

Price $2.00

JOURNAL OF CLINICAL PSYCHOLOGY
5 Pearl Street Brandon, Vermont

EDITORIAL OPINION

DAWN OF A NEW DAY

Perhaps the most significant and exciting discovery of the last 50 years in basic science psychology has been the development of statistical methods of factor analysis. Although factor analysis is still in its infancy with only a relatively small number of statistical experts being proficient in its use, its promise for the future is so great that it may well revolutionize our whole systematic approach to problems in basic and applied science. In spite of many unsolved technical difficulties with current factorial methodology, the accomplishments of the last 20 years have been so impressive as to require a reevaluation of our whole theoretical approach. For the first time, methods have become available for objectively identifying and quantifying psychological factors in terms of the results from P, Q and R correlational techniques. Sufficient research has accumulated so that the basic dimensions of behavior organization have been identified in terms of standard factors defined in operational terms, cross-identified by means of suitable marker variables, and interpreted by committees of competent scientists. The list of all those who have contributed to these important developments is too long to be detailed here, but special commendation must go to such pioneers as Charles Spearman, C. L. Burt, L. L. Thurstone, G. Thomson, K. J. Holzinger, Raymond B. Cattell, J. P. Guilford and William Stephenson. The important thing is for all clinical psychologists to become aware of the tremendous impact which the methods and results of factor analysis must have upon psychological practices in the future.

First, we must reorient our theory in terms of factors already demonstrated. It is no longer valid to couch our thinking in terms of the theoretical abstractions of arm chair psychology. Such general concepts as that of *intelligence, affective life, personality, motivation, perception, memory,* etc., must now be dealt with in terms of component factors operationally defined.

Second, the whole map of psychological space must be redrawn by relocating the relative positions of *general, group* and *specific* factors as they contribute to the organizational composition of personality integration.

Thirdly, research design must be reconstituted to deal with stable factors as they are progressively discovered. It is no longer valid to utilize the crude dimensions of the era of prefactorial theorizing. Statistical and experimental research in the future must be redesigned in order to deal with the finer differentiation of factors now made possible. Factorial findings must be integrated into research design.

Fourth, the education and training of future generations of psychologists must include systematic instruction in the methods of factor analysis so that their entire thinking becomes oriented thereto.

Fifth, the profession as a whole must assume responsibility for establishing periodic scientific surveys of accumulating research in the field in order to reach agreement on the meanings of the data. Ideally, committees of competent psychologists should be appointed by the APA to integrate and interpret the whole research program so as to avoid confusion concerning the meanings of disparate and seemingly conflicting studies. Fortunately, progress has already been made in this direction by technical conferences sponsored by the Educational Testing Service and reported by John W. French.

Sixth, as better understanding of the rationale of factorial techniques is disseminated, it is to be hoped that current inconsistencies of technique will be resolved. Factorial methods must be standardized by group agreement on technical details, research design, standardization of marker variables, and nomenclature so that research from various laboratories becomes comparable.

Seventh, it is desirable that the profession should develop a universal index and nomenclature for factors discovered in order to prevent a terminological chaos. Raymond B. Cattell has already offered an universal index for psychological factors whose principles and application should be understood by every clinical psychologist.

Finally, a clear differentiation and cross-identification should be attempted on data obtained by such different operational methods as life history, introspective reporting and self-ratings, objective behavior ratings, objective tests and projective tests. Until the advent of factor analysis,

it was not properly appreciated that data obtained by these different methods might not be exactly comparable even though labelled and interpreted by common terms. It is only by suitable marker variables that the same factors may be cross-identified in researches utilizing different operational approaches.

It is inevitable that the transition from the pre- to the post-factorial era will be associated with a great deal of confusion and hesitant fumbling as individuals and groups attempt to readjust their thinking. Much of this could be avoided if the APA would organize official committees and groups to facilitate the many cooperative efforts which must be accomplished in order to bring order out of chaos. The AMA has official councils or bureaus on standards, practices, medical services, research, foods and drugs, and many other special problems. On February 7-9, 1955, the AMA held the first official Standard Nomenclature Institute including lectures on Standard Nomenclature systems and coding practices. A great deal of wasted energy could be prevented if the APA would immediately establish suitable machinery for quickly coordinating many of the activities which need to be undertaken in this field which cuts across the interests of many divisions.

F. C. T.

BOOK NOTICES

EATON, J. W. and WEIL, ROBERT J. *Culture and Mental Disorders.* Glencoe, Ill.: Free Press, 1955, pp. 254. $4.00

This is a socio-psychiatric study of Hutterite colonies to discover the influence of such a restricted homogeneous culture upon the incidence of mental disorders. Although this religious sect had previously been considered to be relatively free of neuropsychiatric disorders, this study indicates that the "simple way of life" does not in itself inevitably result in mental health.

LEHNER, G. F. J. and KUBE, E. *The Dynamics of Personal Adjustment.* New York: Prentice-Hall, 1955, pp. 498. $7.00.

This book discusses the psychology of adjustment in the principal situations of all life periods. An attempt is made to show the interaction of personal needs and social forces in the development of behavior. Written in a clear coherent style, it will be suitable as an introductory textbook for students and laymen.

DEUTSCH, F. and MURPHY, W. F. *The Clinical Interview. Vol. II: Therapy.* New York: International Universities Press, 1955, pp. 335. $7.50

The authors define *sector therapy* as involving a goal-limited psychotherapy developed from the associative anamnesis and acting directly and systematically upon a limited chosen sector of the patient's symptoms. It is assumed that therapeutic gains achieved in limited sectors will inevitably affect the whole personality. This book consists of a large number of excerpts from clinical interviews which have been selected for teaching purposes to illustrate the methods used.

BURT, CYRIL. *The Subnormal Mind.* New York: Oxford University Press, 1955, pp. 391. $5.00.

This is the third edition of a standard British textbook on mental deficiency. It provides a general orientation to the nature of mental disabilities.

HILL, L. B. *Psychotherapeutic Intervention in Schizophrenia.* Chicago: University of Chicago Press, 1955, pp. 216. $5.00.

THORNE, F. C. *Principles of Psychological Examining.* Brandon, Vt.: Journal of Clinical Psychology, 1955, pp. 494. $7.50.

A systematic textbook of applied integrative psychology.

BASOWITZ, H., KORCHIN, S. J., PERSKY, H. and GRINKER, R. R. *Anxiety and Stress.* New York: McGraw-Hill, 1955, pp. 320. $8.00.

This is an interdisciplinary study of anxiety and stress in the real life situation of paratroop training. In general, men failing in training did show more stress response, however the situation is very complex, and several types of anxiety may be operating.

BETTELHEIM, B. *Truants from Life.* Glencoe, Ill.: Free Press, 1955, pp. 511. $6.00.

Detailed case histories of the treatment of four severely disturbed children at the Orthogenic School of the University of Chicago.

COHEN, ALBERT K. *Delinquent Boys.* Glencoe, Ill.: Free Press, 1955, pp. 202. $3.50.

U. S. Dept. of Health, Education and Welfare. *Evaluation in Mental Health.* Washington, D. C.: U. S. Gov't Printing Office, 1955, pp. 292. $2.00.

A collection of 984 references to the literature on the evaluation of mental health.

HENRY, GEORGE W. *All the Sexes.* New York: Rinehart, 1955, pp. 599. $7.50.

Dr. Henry is associate professor of clinical psychiatry at the Cornell University Medical College and a recognized authority on homosexuality. This is a popular presentation of the author's long experience with hundreds of homosexuals and bisexuals with illustrative case excerpts and speculative conclusions concerning the relation of psychiatry to public morality in the handling of such people.

DUDYCHA, G. J. *Psychology for Law Enforcement Officers.* Springfield, Ill.: C. C. Thomas, 1955, pp. 404. $7.00.

PODOLSKY, E. *Management of Addictions.* New York: Philosophical Library, 1955, pp. 412. $7.50.

A collection of papers by representative authorities discussing all aspects of alcoholism and drug addiction.

ALDRICH, C. K. *Psychiatry for the Family Physician.* New York: McGraw-Hill, 1955, pp. 276. $5.75.

CRUZE, W. W. *Psychology in Nursing.* New York: McGraw-Hill, 1955, pp. 494. $5.50.

OBERMAYER, M. E. *Psychocutaneous Medicine.* Springfield, Ill.: C. C. Thomas, 1955, pp. 487. $9.75.

NEWTON, NILES. *Maternal Emotions.* New York: Hoeber, 1955, pp. 140. $3.00.

A study of women's feelings toward menstruation, pregnancy, child birth, breast feeding, infant care and other aspects of their femininity.

ABRAMSON, H. A. *Neuropharmacology.* New York: Josiah Macy, Jr. Foundation, 1955, pp. 210. $4.25. Transactions of the first conference.

THE SOUTHARD SCHOOL

Intensive individual psychotherapy in a residential school, for children of elementary school age with emotional and behavior problems.

THE MENNINGER CHILDREN'S CLINIC

Outpatient psychiatric and neurologic evaluation and consultation for infants and children to eighteen years.

Department of Child Psychiatry

THE MENNINGER FOUNDATION

J. Cotter Hirschberg, M. D., Director — Topeka, Kansas; Telephone 3-6494

Principles of Psychological Examining

by FREDERICK C. THORNE, M.D. Ph.D.
Editor, Journal of Clinical Psychology

- The First Comprehensive Textbook of Applied Integrative Psychology.
- Systematic: Presenting a Complete System for the Differential Diagnosis of the Primary Etiologic Factors Which Actually Organize Socially Significant Behaviors.
- Complete: More Extensive Diagnostic Classifications of Etiologic Factors.
- Eclectic: A theoretical system is outlined for relating all pertinent theories and facts in basic science psychology and clinical psychological science.
- Advanced: More than a textbook of psychometrics. Tests and measures are presented only as they relate to more fundamental problems of psychological diagnosis.

PRICE: $7.50

SPECIAL OFFER TO APA MEMBERS $6.00

ORDER NOW

JOURNAL OF CLINICAL PSYCHOLOGY
5 Pearl Street — Brandon, Vermont

d is for diagnosis

D is for Devereux Schools

a score of residential schools and camps on both the East and West Coasts, with unexcelled facilities for both diagnosis and treatment of childhood emotional disorders.

Each Devereux School is a fully equipped and staffed unit on its own, separate campus -- especially designed to meet the needs of a particular age group with common intellectual and emotional characteristics. Between them, they offer a full educational program from pre-kindergarten through junior college, with special provisions for superior, average-normal, and slow-learning youngsters.

The Schools also make full provision for medical and psychiatric therapy -- including formal psychoanalysis, when it is indicated.

You are invited to visit these "special schools" at any time; or if you prefer, further information may be had by writing

JOHN M. BARCLAY
Director of Development
Devereux Schools
Devon, Pennsylvania

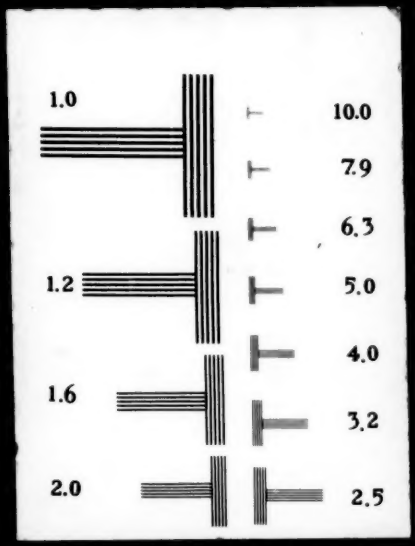
"HUMANI NIHIL A" "ME ALIENUM PUTO"
THE DEVEREUX FOUNDATION
CALIFORNIA PENNSYLVANIA

Devereux Schools

UNDER THE DEVEREUX FOUNDATION
Helena T. Devereux, Director

Santa Barbara, California **Devon, Pennsylvania**

RESOLUTION CHART



100 MILLIMETERS

INSTRUCTIONS Resolution is expressed in terms of the lines per millimeter recorded by a particular film under specified conditions. Numerals in chart indicate the number of lines per millimeter in adjacent "T-shaped" groupings.

In microfilming, it is necessary to determine the reduction ratio and multiply the number of lines in the chart by this value to find the number of lines recorded by the film. As an aid in determining the reduction ratio, the line above is 100 millimeters in length. Measuring this line in the film image and dividing the length into 100 gives the reduction ratio. Example: the line is 20 mm. long in the film image, and $100/20 = 5$.

Examine "T-shaped" line groupings in the film with microscope, and note the number adjacent to finest lines recorded sharply and distinctly. Multiply this number by the reduction factor to obtain resolving power in lines per millimeter. Example: 7.9 group of lines is clearly recorded while lines in the 10.0 group are not distinctly separated. Reduction ratio is 5, and $7.9 \times 5 = 39.5$ lines per millimeter recorded satisfactorily. $10.0 \times 5 = 50$ lines per millimeter which are not recorded satisfactorily. Under the particular conditions, maximum resolution is between 39.5 and 50 lines per millimeter.

Resolution, as measured on the film, is a test of the entire photographic system, including lens, exposure, processing, and other factors. These rarely utilize maximum resolution of the film. Vibrations during exposure, lack of critical focus, and exposures yielding very dense negatives are to be avoided.